

الدكتور عبد الحميد عبد المجيد البلداوي

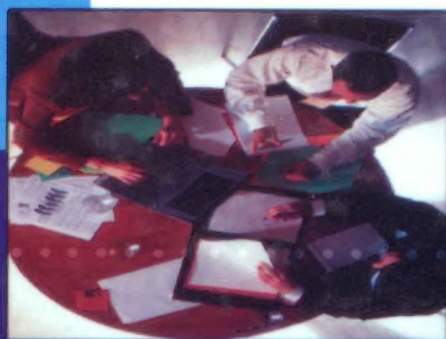
أساليب البحث العلمي والتحليل الإحصائي

التخطيط للبحث وجمع وتحليل البيانات يدوياً

وباستخدام برنامج **SPSS**

METHODS OF SCIENTIFIC RESEARCH & STATISTICAL ANALYSIS

USING SPSS



رقم الإيداع لدى دائرة المكتبة الوطنية
(2007/7/1978)

519.50285

البلداوي، عبد الحميد عبد المجيد
أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع وتحليل البيانات
يدويًا وباستخدام SPSS / عبد الحميد عبد المجيد البلداوي - عمان: دار الشروق، 2007
(240 ص)
ر.إ.: 2007/7/1978
الواصفات: الإحصاء الوصفي//الحواسيب//البحوث العلمية//كتابة البحوث/

• تم إعداد بيانات الفهرسة الأولية من قبل دائرة المكتبة الوطنية

(ردمك) 0 - 318 - 00 - 9957 - 978 ISBN

- أساليب البحث العلمي والتحليل الإحصائي: التخطيط للبحث وجمع وتحليل البيانات يدويًا وباستخدام SPSS .
- تأليف: الدكتور عبد الحميد عبد المجيد البلداوي .
- الطبعة العربية الأولى: الإصدار الثالث 2007 .
- جميع الحقوق محفوظة © .



دار الشروق للنشر والتوزيع

هاتف: 4618190 / 4618191 / 4624321 فاكس: 4610065

ص.ب: 926463 الرمز البريدي: 11118 عمان - الأردن

Email: shorokjo@nol.com.jo

دار الشروق للنشر والتوزيع

رام الله - المصيون: نهاية شارع مستشفى رام الله

هاتف: 2975632 - 2991614 - 2975633 فاكس: 02/2965319

Email: shorokpr@palnet.com

جميع الحقوق محفوظة، لا يسمح بإعادة إصدار هذا الكتاب أو تخزينه في نطاق استعادة المعلومات أو نقله أو
إستنساخه بأي شكل من الأشكال دون إذن خطي مسبق من الناشر.

All rights reserved. No Part of this book may be reproduced, or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without the prior permission in writing of the publisher.

الاعخراج الداخلي وتصميم الغلاف وفرز الألوان والأفلام:

دائرة الإنتاج / دار الشروق للنشر والتوزيع

هاتف: 4618190/1 فاكس: 4610065 ص.ب. 926463 عمان (11118) الأردن

المحتويات

مقدمة..... 13

الفصل الاول

مستلزمات وخطوات تصميم البحث العلمي

- 1-1- تحديد اهداف البحث 17
- 2-1- تحديد مجتمع البحث 18
- 3-1- تحديد وحدة مجتمع البحث 18
- 4-1- تحديد نطاق البيانات المراد جمعها 19
- 5-1- اطار مجتمع البحث 20
- 6-1- تحديد منهجية وطرق التحليل 21
- 7-1- تصميم الاستبانة (الاستمارة) 22
 - (1) مفهوم واهمية الاستبيان 22
 - (2) القواعد العامة لتصميم الاستبيان 22
 - (3) شروط صياغة أسئلة الاستبيان 23
 - (4) أجزاء الاستبيان 25
 - (5) المفاهيم والتصانيف الاحصائية 27
- 8-1- طرق جمع البيانات 35
 - (1) طريقة المشاهدة 35
 - (2) طريقة التسجيل الذاتي 36
 - (3) طريقة المقابلة الشخصية 37

- 38..... (4) طريقة الهاتف
- 39..... (5) طريقة المناقشات العامة
- 40..... 9-1 - اختيار وتدريب العاملين في جمع البيانات (حالة البحوث الكبيرة)
- 41..... 10-1 - المسح التجريبي
- 42..... 11-1 - تعيين التوقيت الزمني الملائم لجمع البيانات
- 43..... 12-1 - آلية العمل الميداني (حالة البحوث الكبيرة)
- 43..... 13-1 - تجهيز البيانات واستخراج النتائج
- 44..... تمارين الفصل الاول

الفصل الثاني

تصميم العينة

- 49..... 1-2 - مقدمة
- 49..... (1) المسوحات الشاملة (التعدادات)
- 50..... (2) المسح بالعينة
- 52..... 2-2 - إجراءات تصميم العينة
- 52..... 3-2 - تحديد حجم العينة
- 59..... 4-2 - أنواع العينات
- 59..... أولاً : العينات العشوائية
- 59..... (1) العينة العشوائية البسيطة
- 65..... (2) العينة العشوائية الطبقية
- 69..... (3) العينة العشوائية المنتظمة
- 71..... (4) العينة العشوائية العنقودية

| | |
|---------|--------------------------------|
| 72..... | ثانيا : العينات غير العشوائية |
| 72..... | (1) العينة المتعمدة (التحكمية) |
| 73..... | (2) العينة الحصصية |
| 74..... | تمارين الفصل الثاني |

الفصل الثالث

تبويب وعرض البيانات

| | |
|-----------|--|
| 79..... | 3-1- مقدمة |
| 80..... | 3-2- ادخال البيانات باستخدام برنامج SPSS |
| 84..... | 3-3- التوزيع التكراري Frequency باستخدام SPSS |
| 93..... | 3-4- التوزيع التكراري المتعدد Cross tab باستخدام SPSS |
| 97..... | 3-5- تفسير مخرجات Cross tabs |
| 97..... | 3-6- توزيع التكرارات على فئات باستخدام برنامج EXCEL |
| 100.... | 3-7- الرسوم والاشكال البيانية باستخدام برنامج SPSS و EXCEL |
| 101..... | اولا : باستخدام برنامج SPSS |
| 103..... | ثانيا : باستخدام برنامج EXCEL |
| 105 | (1) المنحنيات والخطوط البيانية التكرارية والمتجمعة |
| 110..... | (2) الاعمدة البيانية |
| 113..... | (3) الدائرة البيانية |
| 114..... | (4) الرسوم والصور البيانية |
| 115..... | 3-8- الطريقة اليدوية في تبويب وعرض البيانات |
| 115..... | (1) التوزيع التكراري البسيط |

- 120..... (2) التوزيع التكراري المتجمع
- 122..... (3) التوزيع التكراري المزدوج
- 124..... (4) التوزيعات النوعية (الوصفية) والزمنية والجغرافية
- 124..... (5) العرض البياني
- 127..... تمارين الفصل الثالث

الفصل الرابع

مقاييس النزعة المركزية (المتوسطات) و التشتت

- 131..... 1-4- مقدمة
- 131..... 2-4- استخدام الحاسوب مع برنامج SPSS
- 132..... 3-4- الطريقة اليدوية
- 132..... (1) الوسط الحسابي
- 139..... (2) الوسيط
- 144..... (3) المنوال
- 148..... (4) العلاقة التقريبية بين الوسط الحسابي والوسيط والمنوال
- 149..... (5) الوسط الهندسي
- 152..... (6) الوسط التوافقي
- 155..... 4-4- مقاييس التشتت (التباين)
- 155..... (1) المدى
- 156..... (2) الانحراف المعياري
- 159..... 5-4- مقاييس التماثل والالتواء
- 161..... تمارين الفصل الرابع

الفصل الخامس الارتباط

| | |
|----------|-------------------------------------|
| 165..... | 1-5- مقدمة |
| 165..... | 2-5- استخدام الحاسوب مع برنامج SPSS |
| 169..... | 3-5- الطريقة اليدوية |
| 169..... | (1) معامل الارتباط البسيط r |
| 171..... | (2) معامل الارتباط المتعدد R |
| 172..... | (3) معامل الارتباط الجزئي |
| 174..... | (4) معامل ارتباط الرتب |
| 176..... | (5) معامل الاقتتران |
| 177..... | (6) معامل التوافق |
| 179..... | تمارين الفصل الخامس |

(الفصل السادس)

التحليل باستخدام الطرق متعددة المتغيرات

| | |
|----------|--|
| 183..... | 1-6- تحليل الانحدار Regression Analysis |
| 183..... | (1) مقدمة |
| 184..... | (2) استخدام الحاسوب مع برنامج SPSS |
| 184..... | اولا: اجراءات مدخلات تحليل الانحدار |
| 190..... | ثانيا: تفسير مخرجات تحليل الانحدار |
| 192..... | 2-6- تحليل المركبات Principal Component Analysis |

| | |
|----------|--|
| 192..... | (1) مقدمة |
| 193..... | (2) اجراءات مدخلات تحليل المركبات |
| 198..... | (3) تفسير مخرجات تحليل المركبات |
| 198..... | 3-6- الطريقة اليدوية في تحليل الانحدار الخطي |
| 198..... | (1) مقدمة |
| 201..... | (2) استخدام نموذج الانحدار للتنبؤ |
| 202..... | تمارين الفصل السادس |

الفصل السابع

اختبار الفروض وتحليل التباين

| | |
|----------|---|
| 205..... | 7-1- مقدمة |
| 205..... | (1) الفروض Hypotheses Testing |
| | (2) الخطأ من النوع الاول والخطأ من النوع الثاني |
| 205..... | Type I & II Error |
| | (3) اختبار من جانب واحد واختبار من جانبيين |
| 206..... | One & Two Sides Test |
| 207..... | 7-2- استخدام الحاسوب مع برنامج SPSS |
| 207..... | (1) الاختبار الاحادي One Sample T-test |
| 207..... | اولا: المفهوم والمدخلات |
| 209..... | ثانيا: تفسير مخرجات الاختبار الأحادي |
| | (2) الاختبار في حالة عدم تساوي التباين (مجتمعين مستقلتين) |
| 209..... | Two Independent Samples |
| 209..... | اولا: المفهوم والمدخلات |

| | |
|-----------|--|
| 212..... | ثانيا: تفسير مخرجات اختبار عينتين مستقلتين |
| 213..... | (3) اختبار T المقارنات الزوجية Paired Data T-test |
| 213..... | اولا: المفهوم والمدخلات |
| 215..... | ثانيا: تفسير مخرجات استخدام T-test للمقارنات الزوجية |
| 216..... | (4) اختبار مربعات كاي Chi Square test |
| 216..... | اولا: المفهوم والمدخلات |
| 219..... | ثانيا: تفسير مخرجات استخدام اختبار مربعات كاي |
| 220 | (5) تحليل التباين بمعيار واحد One-Way Analysis of Variance |
| 220..... | اولا: المفهوم والمدخلات |
| 225..... | ثانيا: تفسير مخرجات تحليل التباين بمعيار واحد |
| 226..... | 3-3- الطريقة اليدوية في إجراء الاختبارات وتحليل التباين |
| 226..... | (1) الاختبار الاحادي One Sample T test |
| 227 | (2) الاختبار مجتمعين مستقلين Two Independent Samples test |
| 227..... | (3) اختبار المقارنات الزوجية Paired Data T Test |
| 228..... | (4) اختبار مربعات كاي Chi Square Test |
| 229..... | (5) تحليل التباين بمعيار واحد One Way Analysis of Variance |
| 231..... | تمارين الفصل السابع |
| 234..... | الملاحق |
| 240..... | المصادر |

مقدمة

ان الذي لا اختلاف عليه هو اهمية البحوث والدراسات في عصر المنافسة وتحليل الكلفة والعائد والحلول السريعة والناجعة للظواهر الاجتماعية والاقتصادية، والسرعة في عملية التطوير والابتكار من اجل المواكبة والبقاء. الا ان المهم هو ان تكون هذه البحوث قائمة على الدقة العالية والموضوعية العلمية الرصينة والحصول عادي نتائجها باقل كلفة واقصى سرعة ومردود.

ان انجاز بحوث كهكذا مواصفات وخصائص لابد وأن تستند على التحليل العلمي الذي يعتمد الاساليب والطرق الاحصائية الكفوة المعززة بمعايير ومقاييس كمية وعلمية عالية المعنوية. ان مثل هذه الاساليب العلمية الكفوة هي ليست صعبة المنال بل في متناول الجميع بكل سهولة ويسر، لكن المهم في الامر هو التوجه الى استخدامها، والالمام في تفسير مخرجاتها، واخيرا حسن اختيار الاسلوب الذي يناسب الحالة التي تحت البحث والدراسة، وهي متطلبات متواضعة، كل ما تحتاجه هو الجدية والرغبة للباحث او الدارس. ومن بين المتوفر واغلبنا في حاجة اليه في العمل البحثي هو برنامج SPSS وبرنامج EXCEL وغيرها الكثير، الا ان الاول هو الاكثر اهمية للباحثين عموما لما تؤول اليه نتائج من عمق وتفاصيل تفي بحاجة غالبية بحوث ودراسات اليوم.

اما الامر المهم الاخر لاي بحث ودراسة فهو حسن التهيئة والتحضير في جمع البيانات والمعلومات الاحصائية التي ستخضع لعملية التحليل، وفي حسن اختيار العينة التي ستجمع منها هذه البيانات والمعلومات، لانعكاس ذلك على مصداقية واعتمادية النتائج التي يتوصل اليه الباحث.

لقد تم وضع هذا الكتاب نصب عينيه تغطية المستطاع من هذه الاولويات في العمل البحثي على الاقل من خلال تناول الطرق الاكثر تكرارا في الاستخدام من قبل عموم الدارسين والباحثين، فقد تم البدء بالطرق لمستلزمات التهيئة والتحضير بتفصيل نسبي مفيد، تلا ذلك سرد اسلوب تصميم عينة البحث وفق الاسس الاحتمالية

العشوائية وبتبسيط وتركيز على الجوانب التطبيقية وبعيدا عن تعقيدات المفاصل النظرية. وبدأ الفصل الثالث في تناول عملية تبويب وعرض البيانات التي تم جمعها ميدانيا وكيفية تهيئتها لأغراض استخدام الحاسوب وبالتحديد لبرنامج SPSS، والتطرق بذات الوقت الى صيغ ومعادلات وخطوات استخدام هذه الطرق في فقرات التحليل اليدوي بغية التعرف على الاسس التي تتجز بواسطة عملية التحليل عند استخدام الحاسوب. وفي الفصل الرابع تم تناول مقاييس النزعة المركزية (المتوسطات) والتشتت، وكما في جميع الفصول بكلا الحالتين ايضا، حالة استخدام الحاسوب وبدونه، ثم تناولنا الارتباط بمختلف انواعه في الفصل الخامس، وتم تخصيص الفصل السادس للأساليب متعددة المتغيرات. وتم فيه التطرق لكل من الانحدار وتحليل العوامل (المركبات) كنماذج لهذه الأساليب الاحصائية وباعتبارها الأكثر استخداما واهمية في العمل البحثي، وكان اختبار الفروض وتحليل التباين هما موضوع الفصل السابع الذي شمل مقدمة تمهيدية عن مفهوم وحالة استخدام كل نوع من الاختبارات و الفرضية التي يقوم عليها.

مع الإشارة الى ان الاجمال والاختصار ان حصل في بعض المواضيع فهو يعود لسببين، الاول بغية عدم ارباك الطلبة واغلب الباحثين بالتفاصيل وبتعدد الأساليب التي قد يحتاج بعضها الى اسس نظرية قد تحتمل الصعوبة والتعقيد والوقت عند التعامل معها، والسبب الثاني هو لكي لا يدعوا كبر حجم الكتاب الى الشكوى المستمرة من قبل الطلبة وغيرهم من صعوبة حمله كما حصل للمؤلف في مطبوعات سابقة.

أملاً ان يحقق الكتاب الفائدة للدارسين والباحثين داعيا للجميع بالتوفيق، والله هو ولي التوفيق. والحمد والشكر لله رب العالمين.

المؤلف



مستلزمات و خطوات تصميم البحث العلمي
PHASES & REQUIRMENTS OF SCIENTIFIC
RESEARCH DESIGN

1-1- تحديد أهداف البحث Research Objectives

إن الخطوة الأولى والأساسية لأي بحث أو دراسة هي تحديد أغراضها أو الأهداف المتوخى الوصول إليها، بما في ذلك الفرض أو الفروض المطلوب اختبارها، لاجل تحديد مصدر المعطيات (البيانات) وطبيعة ونوعية وشمولية هذه المعطيات. لذا لا بد من أن يكون الهدف (أو الأهداف) تتسم بالشفافية والوضوح وعلى درجة معقولة من التفصيل لتكون على علم كاف بالمعطيات اللازم تغطيتها. فمثلاً إذا كان هدف البحث هو دراسة " مستوى خدمات النقل العام "، عندها يجب أن نوضح وبالتفصيل إذا كان الأمر سيقصر على وسائل النقل فقط أم أن ذلك يتضمن تطوير شبكات الطرق وتحسين الخدمات المرتبطة بعملية النقل وهكذا. ولو تناولنا مثلاً آخر، وليكن دراسة عن قطاع الصناعة فلا بد من معرفة أن كنا بصدد التوصل إلى مستوى الصناعة المستهدفة من ناحية جودتها إدارياً وإنتاجياً، أم الهدف هو لتوفير بناء مؤشرات الحسابات القومية والتعرف على المشاكل التي يواجهها القطاع الصناعي، أو على فرص الاستثمار المتاحة في هذا القطاع وقد يكون الأمر يتعلق بواحد أو أكثر من الأهداف التالية التالية:

1. التعرف على أنواع الصناعات الاستخراجية والتحويلية المختلفة الموجودة، وتوزيعها الجغرافي، وحجم إنتاج كل منها.
2. التعرف على كميات وقيم مستلزمات الإنتاج الصناعي حسب أنواعها ومصادرها.
3. التعرف على منافذ توزيع المنتجات الصناعية (السوق المحلي، التصدير).
4. تقدير حجم ومصدر رأس المال المستثمر في القطاع (وطني، عربي، أجنبي).
5. تقدير حجم العمالة حسب النوع والجنسية والقطاع والكيان القانوني والنشاط الاقتصادي والمهنة والأجور والرواتب.
6. التعرف على الطاقة الإنتاجية المستغلة والمعطلة، وأسباب التعطيل.

وعادة ما يتم نقل هذه الأهداف إلى صيغة جداول، تدعى بجداول الإنتاج (المخرجات Output)، والتي يراعى في تصميمها طبيعة العلاقات الإحصائية المستهدفة بين المتغيرات (Variables) ذاتها أو بين المتغيرات ووحدات الملاحظة

(Observations) التي قد تكون المنشآت أو المناطق الجغرافية أو غيرها، ليتم في ضوءها تصميم الاستبانة (الاستمارة الاحصائية) التي سيلبي التطرق إليها لاحقاً في هذا الفصل .

2-1 - تحديد مجتمع البحث . Population Scope

بعد تحديد الهدف (أو الأهداف)، يتطلب الأمر تحديد المجتمع المشمول بالبحث الذي سنقوم بجمع المعطيات منه، مع ضرورة معرفة حدوده، وحدود احتياجنا منه. فبالنسبة للمثال الأول الوارد في الفقرة (1-1) أعلاه، ينبغي أن نحدد في هذه المرحلة إن كان المقصود هو اخذ عينة من كافة مجتمع النقل العام (مسافرين وبضائع) في الدولة، أم من نقل المسافرين فقط، وإن كان عن نقل المسافرين فهل المقصود من داخل المدن، أم النقل بين المدن، أم من كليهما، وهل سيضم كل المناطق، أم من مناطق محددة، فإن كان المقصود مناطق محددة عندها يجب تسمية تلك المناطق، مع تسمية المجالات الأخرى التي يقتضي تغطيتها. أما على نطاق المثال الآخر المتعلق بالمسح الصناعي، فيتم توضيح إن كان المسح سيغطي المنشآت العاملة في كافة المحافظات (حضر وريف) التي تمارس أنشطة الصناعات الاستخراجية، والصناعات التحويلية غير البترولية.... الخ. أم إن الأمر سيقصر على صناعة محددة وفي محافظة معينة وإلى غير ذلك.

3-1 - تحديد وحدة مجتمع البحث Population Observation

أن تحديد مفهوم وحدة المجتمع (Observation) التي ستجمع منها معطيات الدراسة هو أمر في غاية الأهمية لإجراء المقارنات الجغرافية والزمنية وغيرها، لذا من الواجب تحديدها من غير التباس أو غموض بحيث تكون واضحة التعريف، سهلة التعيين والعد. فمثلاً إذا كانت الوحدة المستهدفة في البحث هي الأسرة، كان لزاماً علينا التعريف الدقيق لمفهوم الأسرة، هل تعني الأب والأم والأولاد، أو تعني كل من يسكن مع الأب والأم والأولاد من أقرباء، أم أنها تعني كل من يسهم في نفقات الأسرة ودخلها سواء أكان هؤلاء من الأقرباء أو من غيرهم، وإن كانت الوحدة

هي المشروع الصناعي فهل يشمل ذلك الصناعات الكبيرة والصغيرة ام صناعات لايتجاوز رأسمالها حدًا معيناً... الخ. بكلمة أخرى ينبغي ألا يترك مفهوم وحدة العد مبهما أو خاضعا للاجتهاد الشخصي بل يجب أن نعرف مسبقا ماهية الوحدة المشمولة، مراعين في ذلك المفاهيم والتصانيف الدولية والمحلية المقررة رسميا والتي سيلبي الاشارة إليها لاحقا. وبعبكسه سنأتي المعطيات التي يتم جمعها مضلله عند إجراء المقارنات الدولية أو الزمنية أو الجغرافية ونتائج تحليلها غير معبرة عن الواقع.

4-1 - تحديد نطاق البيانات المراد جمعها Scope of Data

ينبغي أن تكون المعطيات التي تجمع من العينة ذات علاقة مباشرة بالهدف من الدراسة، مما يستوجب تحديدها لئلا نهمل معطيات أساسية، أو نزيد من معطيات ليست لها علاقة بأهداف الدراسة الممثلة بجداول الإنتاج التي تعبر أيضا عن طبيعة العلاقات الإحصائية كالجغرافية والزمنية والديموغرافية والاقتصادية وغيرها، ولتجنب التكلفة غير المبررة. وليستعان في ضوءها بعد ذلك بتصميم استمارة (استبيان) البحث. وفي ضوء ما تقدم لو تأملنا بمثالنا الوارد في الفقرة (1-2) والمتعلق بدراسة تخطيط النقل العام وتطويره، نجد أننا بحاجة إلى تحديد ماهية المعطيات التي تقي بالهدف وتغطي حاجة البحث أو الدراسة، فجانبا تطوير وسائط النقل يعني تغطية خصائص المسافرين من مستخدمي هذه الوسائط، ويتمثل ذلك بالدخل والعمر والنوع والمهنة والغرض من الرحلة (إن كانت رحلة عمل أو رحلة غير عمل)، وكذلك استطلاع رغباتهم وآرائهم بشأن خصائص واسطة النقل التي يرغبون فيها من ناحية سعتها (عدد المقاعد) ودرجة الأمان فيها وسرعتها، ومستوى الأجور والمجال المخصص للحقائب والعفش ومستوى الراحة والملاءمة وإلى غير ذلك. ثم نتناول الجانب الثاني المتعلق بشبكة الطرق إن كان ضمن الاهداف المطلوبة ونحدد المطلوب من المعطيات لدراسة هذا الجانب، وقد نجد بان حاجة الدراسة هي معطيات تتعلق بحجم حركة المرور على الطرق ومنشأ (Origin) ومستقر (Destination) هذه الحركة وأنواع وسائط النقل المستخدمة (صالون، بيك أب، لوري، حافلة، شاحنة، عجلة زراعية وغيرها) موزعة حسب ساعات اليوم. نقوم بعد ذلك بدراسة حاجة الدراسة إلى

المعطيات المتعلقة بتطوير خدمات محطات النقل، وذلك باستطلاع آراء المسافرين عن طبيعة الخدمات التي يرونها مناسبة لتوفيرها في هذه المحطات، من أماكن استراحة وانتظار ومكاتب حجز وأماكن بيع صحف ومجلات ومطاعم وأسواق بيع سلع خفيفة وهدايا وتوفير هواتف عمومية وما إلى ذلك .

1-5- إطار مجتمع البحث. Research Population Frame

و الإطار عبارة عن وصف لما هو متوافر من معطيات عن مفردات المجتمع المطلوب دراسته والذي ستسحب منه العينة. وعادة ما يعتمد في توفير هذه المعطيات على نتائج المسوحات الإحصائية الشاملة أو ما هو متوفر في سجلات الجهات الرسمية المختصة، كأساس لتكوين الأطر، وقد تجرى عمليات تحديث على هذه الأطر في حالة مضي زمن عليها، وقد تتخذ الأطر شكل خارطة تضم المواقع المطلوب بحثها كالمقاطعات أو القرى أو المزارع أو مواقع المصانع، أو شكل قوائم بأسماء وعناوين مفردات المجتمع، فإذا كانت المفردة الإحصائية هي المصنع مثلا فإن الإطار يصبح عبارة عن قائمة تضم أسماء المصانع في منطقة الدراسة وعناوينها. وفي كثير من الحالات يتم الاعتماد لهذا الغرض على القوائم التي تضم أسماء المستفيدين من خدمات الكهرباء والماء المتوافرة لدى المؤسسات أو الجهات الرسمية المعنية بهذه الخدمات. وفضلا عن كون الأطر هي من مستلزمات تصميم العينة، فإنها تعتمد أيضا لأغراض إدارة المسح وتنفيذه، إلى جانب تنظيم العمل الميداني من خلال ما توفره من معطيات تفيد في تحديد المواقع المشمولة بالمسح عند التحاق الباحثين والمشرفين بمواقع عملهم عند متابعة العمل الميداني. فعلى سبيل المثال عند ذكر إطار المستشفيات يعني توفر معطيات عن جميع المستشفيات في الدولة في شكل قوائم بأسمائها وعناوينها واختصاصاتها وعدد الأسرة فيها وإلى غير ذلك. لابد من الأخذ بنظر الاعتبار التغيرات التي حصلت على معلومات الإطار الذي يعتمد لسحب العينة لكي تأتي العينة ممثلة لخصائص المجتمع من جهة ولأغراض تخطيط إدارة العمل الميداني من جهة أخرى، وان مراعاة عملية التغيرات هذه هي ما تدعى "تحديث الإطار"، ولأجل ذلك تم إضافة مثلا المنشآت الصناعية

التي استحدثت بعد تاريخ سجلات الجهات المعنية من خلال معرفة التراخيص الممنوحة بعد التاريخ الذي يعود اليه الاطار. ويمكن الاستعانة في حالة هذا المثال بغرفة التجارة او الصناعة أو المؤسسة العامة للصناعة او غيرها، وبصورة عامة، فلكي يكون الإطار صالحا ينبغي أن تتوافر فيه الشروط التالية:

1. أن يكون حديثا ويعود لتاريخ قريب من الزمن الذي تؤخذ منه العينة.
2. أن يحتوي على جميع مفردات المجتمع المراد دراسته.
3. أن لا يحصل تداخل بين مفردات المجتمع (أي عدم حصول تكرار في ظهور أي من الوحدات).

6-1- تحديد منهجية وطرق التحليل

Methodology and Methods of Analysis

إن الأهداف التي تتوخاها أي دراسة يمكن تحقيقها باعتماد منهجيات وطرق تحليل مختلفة، وإن اختيار ما هو مناسب من بينها يرتبط بظروف الدراسة من إمكانيات فنية ومالية وبشرية. وغالبا ما يكون لكل منهجية طرقها التحليلية التي قد تستلزم في بعض جوانبها حاجة مختلفة في طبيعة المعطيات وفي مستوى تفصيلها، فمثلا لو تأملنا بمثال دراسة تطوير النقل العام لوجدنا أن بالإمكان اعتماد أحد نوعين أو أكثر من المنهجيات. فهناك ما يدعى بالمنهجية التقليدية التي تتطلب معطيات تجميعية Aggregate data وتكون على مستوى مناطق جغرافية Zones وهذه المنهجية تحتاج إلى تفاصيل وإلى عينة كبيرة نسبيا. أما النوع الآخر الذي يدعى بالمنهجية السلوكية Behavioral methodology التي تحتاج إلى استخدام طرق ونماذج احتمالية، كطريقة لوجت Logit method أو طريقة Probit أو طريقة الانحدار، ولا تحتاج هذه المنهجية أكثر من مرحلة تحليلية واحدة يتم فيها تحديد حصص كل من وسائط النقل المتوفرة من إجمالي حجم الطلب، من خلال استخدام النماذج التحليلية المذكورة وتضمنينها متغيرات تتعلق بخصائص المسافرين ووسائط وطبيعة الرحلات المتحققة وخصائص الطريق، وإن المعطيات التي تحتاجها تكون على مستوى المفردة Disaggregate level، ويكتفى بعينة واحدة صغيرة نسبيا.

1-7 - تصميم الاستبانة (الاستمارة) Questionnaire Design

1. مفهوم وأهمية الاستبيان Questionnaire Definition

الاستبيان الإحصائي عبارة عن صحيفة أو كشف يتضمن عددا من الأسئلة تتصل باستطلاع الرأي أو بخصائص أية ظاهرة متعلقة بنشاط اقتصادي أو اجتماعي أو فني أو ثقافي. ومن مجموع الإجابات عن الأسئلة نحصل على المعطيات الإحصائية التي نحن بصدد جمعها. إن لتصميم الاستبيان والأسئلة التي يتكون منها تأثيراً مباشراً على نوعية المعطيات ودرجة دقتها. لذا يحتاج التصميم إلى عناية فائقة وإلمام تام بحالة المشمولين بالمسح الإحصائي وفهم لتقاليدهم وأمورهم الاقتصادية والاجتماعية، وحتى لمذلولات الألفاظ واللغة المتداولة بينهم. ومن الجدير بالذكر ان تصميم الاستبيان يأتي بعد الانتهاء من تحديد طبيعة المعطيات الإحصائية المطلوب جمعها، والتي كما ذكرنا يتم تمثيلها بجداول إنتاج تعرض الصيغة النهائية للمعطيات ولطبيعة العلاقات الإحصائية بين المتغيرات المستهدفة.

2. القواعد العامة لتصميم الاستبيان

Questionnaire Design Rules

أولاً: ينبغي ان يكون حجم الاستبيان مناسباً، ونوع الورق المستعمل يتحمل الكتابة، ويكون لونه مقبولا، وتكون الطباعة جيدة وسهلة القراءة، وإذا كان الاستبيان مكوناً من عدة صفحات فإنه يستحسن ان يكون على شكل كراس.

ثانياً: مراعاة التنفيذ الآلي لتبويب المعطيات وتحليلها إذا كان في النية استخدام الحاسب الآلي، وذلك بتخصيص حقول للرموز Coding خاصة للإجابة على كل سؤال، وتكييف الاستبيان بما يتلاءم وهذا الغرض، ويتم أيضاً مراعاة ما إذا كانت عملية الترميز تقع ضمن الإجابة؛ أي الترميز المسبق Pre-coded ام تتم لاحقاً بعد ملء الاستمارة.

ثالثاً: ضرورة أن يضم الاستبيان الحد الأمثل من الأسئلة قدر الإمكان، وان تحقيق ذلك يستوجب مراجعتها عدة مرات للتأكد من خلوها من الأسئلة التي لا تخدم أهداف الدراسة.

3 شروط صياغة أسئلة الاستبيان Questions Structure Rules

أولاً: الأخذ بنظر الاعتبار أن الأسئلة موجهة إلى أفراد مختلفين في مستوياتهم ومؤهلاتهم الثقافية والتعليمية وحتى أحيانا في عاداتهم الاجتماعية، مما يستدعي الوضوح في صياغة الأسئلة من خلال استعمال عبارات بسيطة لها معنى مألوف وتعطي في الوقت نفسه المعنى المقصود. فمثلا يختلف مفهوم الشركة أو المشاركة عند سكان البادية عنه عند سكان المدن، ففي المدن يرتبط مفهومها بمساهمة مجموعة من الأشخاص برأسمال معين لأجل مزاولة نشاط اقتصادي أو تجاري، في حين يرتبط مفهوم الشركة في البادية بالمشاركة في قطيع الماشية أو الجمال وفي حصص الإنتاج الحيواني.

ثانياً: تجنب الأسئلة الغامضة، كأن يسأل المسافر مثلاً فيما إذا كان مستوى النقل العام في هذه السنة أفضل من مستواه للعام الماضي، مما يجعل الإجابة صعبة ومعقدة لعدم توضيح معنى المستوى وعدم تحديد المعيار المعتمد للقياس، أهو معيار سرعة الوسطة أو معيار الراحة والملاءمة أم دقة مواعيد السفر أم معيار آخر. بالإضافة إلى إمكانية تحديد عدد من المستويات ليقوم المبحوث بتأشير المستوى المناسب لقناعاته أو اعتقاده وهكذا.

ثالثاً: أن تصاغ الأسئلة بحيث تكون الإجابة عليها قاطعة، كأن تكون عبارة عن رقم أو كلمة نعم أو لا أو استخدام إشارات معينة. كما ويفضل ألا تكون الأسئلة من النوع المفتوح، بل حصر جميع الإجابات المحتملة عن كل سؤال وكتابتها أمام السؤال، ليقوم المبحوث بوضع علامة على الإجابة المناسبة، كما هو الحال مثلاً في ذكر المستويات التعليمية عند السؤال عن مستوى التحصيل الدراسي، وكتابة عدد من الهوايات الرئيسية عند سؤال المبحوث عن هوايته المفضلة ليحدد أحدها، فإن لم تكن هوايته بين الهوايات المثبتة فإنها تدخل في فقرة أخرى ينبغي إضافتها إلى فقرة الهوايات المحددة. وذلك بغية التبسيط واختصار الوقت ووضوح المعنى مما يزيد من الدقة، بالإضافة إلى تحقيق هدف التوحيد عند التبويب.

رابعاً: ضرورة ترتيب الأسئلة ترتيباً منطقياً يراعي العلاقة فيما بينها، ويمكن أن يتم ذلك بتقسيم الأسئلة إلى مجموعات متجانسة تحمل عناوين فرعية، مراعين البدء بالأسئلة السهلة التي لا تحتاج إلى تفكير، كذلك المتعلقة بخصائص الشخص كالاسم والعنوان والجنس والعمر وما شابه.

خامساً: ألا تكون الأسئلة من النوع الإيحائي، أي التي توحى إلى المبحوث بإجابات معينة، فلا يسأل مثلاً: هل أنت متدين؟ لأنه ليس من المنتظر أن تكون الإجابة بالنفي، ولكن يكتفي بالسؤال عما إذا كان المشمول يؤدي بعض الشعائر الدينية مثلاً.

سادساً: أن تكون الأسئلة قدر الإمكان بعيدة عن الحساسية أو الإحراج، وألا تعد تدخلا في مسائل شخصية قد تؤدي إلى إزعاج الشخص المبحوث. مع التأكيد أيضا تجنب قدر الإمكان الأسئلة التي تحتاج إلى تفكير وخبرة واسعة.

سابعاً: الابتعاد عن الأسئلة التي تنثير تحيز الشخص، فلا يسأل المبحوث مثلاً هل تأخرت بسبب سوء النقل؟ لأن الإجابة ستكون في الغالب بالإيجاب، فالإشارة إلى السبب تتضمن دلالة الاتفاق عليه وإن لم يكن السبب الرئيسي أو عدمه. أو أن يسأل: هل تشتري الصحف يوميا؟ فقد يدفع حب التفاخر أو الخجل إلى الادعاء بشرائها. وبدلاً من ذلك يمكن مثلاً أن يكون السؤال على النحو الآتي: هل تطلع على الصحف يوميا؟

ثامناً: أما تضمين الأسئلة أكثر من نقطة واحدة، فإذا كان لأحد الأسئلة جزءان فإنه يستحسن أن يكونا سؤالين متتاليين. فلا يسأل مثلاً: هل تمتلك فيديو وتلفزيون؟ فمن الجائز أن يمتلك المبحوث أحدهما فقط. كما لا يستحسن دمج سؤالين معاً مثل: هل تستمع إلى الراديو وأي البرامج تفضل؟

تاسعاً: أن تصاغ الأسئلة بشكل لا يتطلب من المبحوث إجراء عمليات حسابية مطولة أو تستدعي ذاكره حادة ومجهوداً فكرياً، فلا يسأل مثلاً: كم هو عمرك في تاريخ معين، ويكتفى بالسؤال عن تاريخ الميلاد ليقوم الباحث

بعد ذلك بإجراء عملية الطرح لمعرفة العمر. وان لا يسأل مثلا عن معدل عدد الأفراد في الغرفة الواحدة، بل يكفي بالسؤال عن عدد أفراد الأسرة وعن عدد الغرف لمعرفة ذلك.

عاشرا: ضرورة ذكر الوحدات القياسية مثل عدد، كيلو، قدم،... الخ، مع تفضيل المقاييس الكمية والابتعاد قدر المستطاع عن المقاييس الكيفية التي تتوقف على تقدير الشخص المبحوث، فلا يسأل مثلا: هل تذهب إلى المكتبة العامة كثيرا؟ إذ أن (كثيرا) غير محدده ويستحسن تحديد عدد المرات، ليصبح السؤال: اذكر عدد المرات التي تزور فيها المكتبة أسبوعيا؟ أو تحديد المرات على شكل فئات مثل 0-2، 3-5، 6 فأكثر ليقوم المبحوث باختيار أحدها.

أحد عشر: من المفضل إضافة بعض الأسئلة بصيغ مختلفة لا بقصد الإجابة عنها لذاتها وإنما للتأكد من دقة بعض الإجابات الأخرى، كأن يسأل في بداية الاستمارة عن متوسط دخل الفرد الشهري وفي مكان آخر عن متوسط المصروف الشهري لتتم المقارنة بين الإجابتين.

4- أجزاء الاستبيان Questionnaire Parts

بصورة عامة، يتكون الاستبيان من ناحية المحتويات وترتيب الأسئلة من ثلاثة أجزاء رئيسية. أما في الحالة التي يتولى فيها الباحث أو العداد بنفسه أو تحت إشرافه عملية ملء الاستبيان عندها سيتكون من الجزئين الأول والثاني فقط، مع حصول بعض التغيير في الجزء الأول منها. وهذه الأجزاء الثلاثة التي يتكون منها الاستبيان هي:

الجزء الأول: ويتضمن المعطيات المتعلقة باسم الجهة المسؤولة عن المسح الإحصائي وعنوانها. وفي حالة ملء المبحوث ذاته للاستبيان فإن هذا الجزء ينبغي أن يتضمن أيضا مقدمة مكثفة ومعبّره توضح أهمية المسح الإحصائي وأهدافه، مع الإشارة إلى أن المعطيات ستكون سرية واستخدامها سيكون مقتصرًا على الأغراض العلمية فقط. وان مثل هذه المقدمة لا

تظهر لها حاجة عندما يتولى الباحث أو العداد تدوين الإجابات بنفسه، وذلك يعود إلى إمكانية توضيح مثل هذه المقدمة شفويا إلى الشخص المبحوث. كما قد يشتمل هذا الجزء في بعض الحالات على عدد من الأسئلة التي يقوم الباحث بالإجابة عنها من غير الحاجة إلى توجيهها إلى الشخص المبحوث، كذكر اسم المدينة أو المحلة التي يجري فيها المسح، أو ذكر اليوم والتاريخ واسم الباحث واسم مدقق الإجابات وما شابه. فلو افترضنا أننا بصدد جمع معطيات إحصائية لدراسة العوامل المؤثرة في الطلب على النقل بين المدن، لاستخدامها في دراسة لتهيئة المستلزمات المطلوبة من وسائل نقل وأيدي عاملة ومحطات وخدمات وغيرها، وبما يتناسب وحجم الطلب المتوقع، فإن طريقة جمع المعطيات المزمع استخدامها هي التدوين الذاتي (أي يملأ الاستبيان من قبل المبحوث)، وعلى افتراض أن الجهة القائمة بالدراسة هي المؤسسة العامة للنقل، وإن المعطيات سيتم معالجتها يدويا من دون استخدام الحاسب الآلي، فإن هذا الجزء من الاستمارة التي سيتم توزيعا على عينة من المسافرين، سيأخذ الشكل المبين في الصفحة الأولى من النموذج رقم (1-1) المرفق في هذا الفصل .

الجزء الثاني: في هذا الجزء يتم ترتيب الأسئلة الرئيسية المستهدفة في الدراسة، مبتدئين من تلك التي لا تحتاج إلى تفكير كالاسم والجنس والعمر والمهنة وما شابه، على أن يراعى في هذا الترتيب وكما نوهنا سابقا موضوع التجانس بين الأسئلة ومنطقية التسلسل. واستمرارا لمثالنا موضوع الجزء الأول، فإن أسئلة هذا الجزء هي كما هو مبين في الصفحة الثانية من النموذج (1-1)، ومنه يتضح أنه كلما ازدادت الحاجة إلى تفكير للإجابة، فإن موقع السؤال يأتي تسلسله متأخرا، والشئ بنفسه يمكن أن يقال من ناحية ترتيب أقسام هذا الجزء .

الجزء الثالث: أما الجزء الثالث فيتضمن التعليمات الخاصة بشرح الأسئلة وتفسيرها، وفي بعض الحالات عن كيفية ملء الاستمارة، وذلك لكي

تكون مضامين الاستبيان مفهومة للمبحوثين في حالة تولي ملئها بأنفسهم ولكافة العاملين في المسح سواء أكان عملهم ميدانياً أو مكتبياً، وذلك لتلاقي الاجتهادات والتفسيرات الشخصية. وبصورة عامة فمن المفضل ان تكون هذه التعليمات على شكل كراس منفصل في الحالات التي تكون فيها الاستمارة بحاجة إلى شرح واسع لمضامينها. مع التأكيد على ضرورة الاعتماد على المفاهيم والتصانيف الإحصائية المحلية والدولية في وضع هذه التعليمات وذلك كما ذكرنا، للتمكن من اجراء المقارنات الزمنية والمكانية وغيرها.

5- المفاهيم والتصانيف الإحصائية

Statistical Definition & Classifications □

من بين المعايير الأساسية التي يتقرر بموجبها مستوى دقة وفائدة واعتمادية المعطيات الإحصائية لأي مسح إحصائي هو مدى اعتمادها للمفاهيم والتصانيف المقررة رسمياً، لكي تصبح صالحة للمقارنات الدولية والجغرافية والزمنية. كما أن اعتماد الأدلة التي تضم هذه المفاهيم والتصانيف يساعد في توفير الوقت والجهد عند ترميز الاستثمارات وعند تبويب معطياتها. وتهتم المؤسسات والمنظمات الدولية المتخصصة بإعداد هذه الأدلة ودراستها وتوصي باستخدامها بغية التوحيد في مفهوم وموقع البيان الإحصائي. وعادة ما يتم تكيف هذه المفاهيم والتصانيف عند المراتب الفرعية بما يتلاءم وظروف وخصائص كل دولة. وكأمثلة في هذا المجال نستعرض فيما يأتي بصورة مكثفة التصنيف القياسي الموحد للأنشطة الاقتصادية الموحد International Standard Industrial Classification of Economic Activities

التصنيف القياسي الدولي الموحد للأنشطة الاقتصادية:

ويهدف هذا الدليل إلى توفير أطر تساعد وتسهل المقارنات في مجال الإحصاءات الاقتصادية والاجتماعية والخدمات سواء على النطاق المحلي أو الدولي. وبات هذا الدليل يشكل إحدى المستلزمات الفنية الأساسية للعمل الإحصائي في مجال الترميز والتصنيف وإحصاءات التجارة الخارجية وأنظمة الحسابات القومية والعديد من المجالات الأخرى. ورغم أن محاولات مبكرة قد جرت في هذا المجال من قبل عصبة الأمم المتحدة، إلا أن أول تصنيف موحد للأنشطة الاقتصادية

قد ظهر في سنة 1946، وقامت عدد من الدول وجميع المنظمات الدولية بنشر معطياتها الإحصائية وفقاً لهذا الدليل. إلا أن التطور الكبير الذي طرأ في مجال الإحصاء واستخدام التحليل الكمي والنماذج الاقتصادية والتشابك الصناعي وضرورات إيجاد تنسيق بين الدليل والأدلة الأخرى أدت إلى إعادة دراسته ومراجعته وإجراء التعديلات عليه في السنوات 1956 و 1958 و 1968. وكانت المراجعة الثالثة هي آخر ما أنجزه المكتب الإحصائي التابع للأمم المتحدة وذلك في عام 1990. وتتفاوت الفترة التي تقوم بها كل من الدول الأعضاء بالأمم المتحدة بتحديث معلومات الدليل وفقاً لخصائصها، فنجد دولاً لازالت تعتمد المراجعة الأولى وأخرى الثانية وبعضها يعتمد حالياً المراجعة الثالثة (SNA) لسنة 1993 كما هو الحال في دولة الإمارات العربية المتحدة ودول خليجية أخرى حيث قامت بتحديث المراجعة الأخيرة في سنة 1995، فأصبحت عدد الفئات الرئيسية للدليل حالياً 15 فئة بعد أن كانت 9 فئات فقط في المراجعة الأولى، وهذه الفئات هي:

- الزراعة.
- صيد الأسماك.
- التعدين واستغلال الثروات الطبيعية.
- الصناعات التحويلية .
- الكهرباء والغاز والمياه.
- البناء والتشييد.
- التجارة وخدمات الإصلاح (خدمات الصيانة والتصليح).
- الفنادق والمطاعم.
- النقل والتخزين والاتصالات.
- الوساطة المالية.
- العقارات والتأجير وخدمات الأعمال.
- التعليم.
- الصحة والعمل الاجتماعي.
- خدمات المجتمع والخدمات الشخصية الأخرى.
- المنظمات والهيئات الدولية.

فالرمز 5211 مثلاً يشير إلى الفصل (1) من الباب (1) من القسم (52)، حيث تبين الأرقام حسب ترتيبها من اليسار إلى اليمين القسم والباب والفصل. أمّا الأنشطة الاقتصادية فقد أعطيت رموزاً مكونة من ستة أرقام، فمثلاً الرمز (10-3610) يمثل النشاط (10) من الفصل (3610). والجدول رقم (1-1) يعطي نموذجاً للتصنيف الموحد يمثل أقسام وأبواب وفصول فئة الزراعة ونشاط أحد فصوله وفقاً لتجربة دولة الإمارات العربية المتحدة .

جدول (1-1)

فئة الزراعة بموجب التصنيف القياسي الدولي الموحد للأنشطة الاقتصادية

| الفئة | القسم | الباب | الفصل | المسميات |
|-------|-------|-------|-------|---|
| أ | | | | الزراعة |
| | 01 | | | الزراعة والخدمات المتعلقة بها |
| | | 011 | | زراعة المحاصيل والبساتين |
| | | | 0111 | زراعة الحبوب والمحاصيل |
| | | | 0112 | زراعة الخضار ومنتجات المشاتل |
| | | | 0113 | زراعة أشجار الفاكهة والتوابل |
| | | 012 | | تربية الحيوانات وإكثارها |
| | | | 0121 | تربية الماشية والأغنام وإكثارها |
| | | | 0122 | تربية الحيوانات الأخرى |
| | | 014 | 0140 | الخدمات المتعلقة بالإنتاج النباتي والحيواني |

ويدخل ضمن كل فصل عدد من الأنشطة، فعلى سبيل المثال يشمل الفصل 0121 من الباب 012 المتعلق بتربية الحيوانات وإكثارها ما يلي:

| الفصل | النشاط | التفاصيل |
|---------|---|----------|
| 0121 | تربية الماشية والأغنام وإكثارها | |
| 0121-01 | تربية الأبقار (ويشمل المنشآت التي تعمل في تربية الأبقار سواء للاستفادة منها في الأغراض التجارية بإكثارها أو بيعها أو للاستفادة للأغراض الصناعية). | |
| 0121-02 | تربية الأغنام | |
| 0122 | تربية الحيوانات الأخرى | |
| 0122-01 | تربية الدواجن | |
| 0122-02 | تربية النحل وإنتاج العسل | |

مثال (1-1): تعتزم غرفة التجارة والصناعة القيام ببحث عن العاملين في قطاع الصناعة، تتعلق بخصائص هؤلاء العاملين وتوزيعهم الجغرافي، مع إجراء تحليلات إحصائية لأغراض إدارية. والمطلوب هو تصميم استبيان إحصائي للمسح مع مراعاة استخدام الحاسب الآلي في استخراج النتائج، وتوظيف طريقة التدوين الذاتي في جمع المعطيات (أي أن يقوم المبحوثون بملاء الاستبيانات). على وفق التفاصيل التالية:

- 1- النشاط الاقتصادي الذي يعملون فيه وفقاً للتصنيف القياسي الدولي الموحد للأنشطة الاقتصادية ISIC.
- 2- الجنس والعمر (حسب الفئات العمرية التالية): أقل من 18 سنة، 18-39، 40-59، 60 فأكثر.
- 3- الحالة التعليمية وكالاتي: أمي، يقرأ ويكتب، ابتدائية، إعدادية، ثانوية، دبلوم (أعلى من الثانوية وأقل من البكالوريوس)، بكالوريوس، دراسات عالية، أخرى.

الحل (1-1): مبين في نموذج الاستبانة التالي:

غرفة تجارة وصناعة -----

دائرة الدراسات والبحوث

استبانة إحصائية

خاصة للعاملين في الصناعة

أخي المبحوث (أختي المبحوثة):

إن الهدف من هذا البحث هو لتحسين ظروفكم الاقتصادية والاجتماعية والخدمات التي تقدم إليكم، وإن تعاونكم في ملء هذه الاستمارة وإدلاءكم بالمعلومات الدقيقة هو الأساس في تحقيق هذا الهدف، علما بأن استخدام المعلومات سيكون مقتصرًا على الأغراض العلمية، وليس هناك حاجة لذكر الاسم.

وشكرا لتعاونكم

القسم الأول: مكان العمل وطبيعته: حقل خاص بالحاسب

- 1- اسم المؤسسة (شركه أو مصنع) الذي تعمل فيها ()
- 2- عنوان المؤسسة ()
- 3- طبيعة النشاط الاقتصادي للمؤسسة..... ()
- الزراعة----- (01)
- صيد الأسماك ----- (02)
- التعدين واستغلال الثروات الطبيعية ----- (03)
- الصناعات التحويلية----- (04)
- الكهرباء والغاز والمياه----- (05)
- البناء والتشييد----- (06)
- التجارة وخدمات الصيانة والتصليح----- (07)
- الفنادق والمطاعم----- (08)

- النقل والتخزين والاتصالات ----- (09)
- الوساطة المالية ----- (10)
- العقارات والتأجير وخدمات الأعمال ----- (11)
- التعليم ----- (12)
- الصحة والعمل الاجتماعي ----- (13)
- خدمات المجتمع والخدمات الشخصية الأخرى ----- (14)
- المنظمات والهيئات الدولية ----- (15)

القسم الثاني: خصائص المبحوث (أو المبحوثة):

- 1- الجنس: ذكر () أنثى () ()
- 2- العمر: أقل من 18 سنة () 18-39 ()
- 3- الحالة التعليمية: أمي () يقرأ ويكتب () ابتدائية ()
إعدادية () ثانوية () أعلى من الثانوية وأدنى من البكالوريوس
بكالوريوس () دراسات عليا () أخرى () ()

استبانة رقم (2-1)

وزارة النقل

المؤسسة العامة للنقل البري

استبانة إحصائية خاصة

بنقل المسافرين بين المدن

أخي المسافر:

إن الهدف من هذا البحث هو تطوير خدمات النقل وتحسينها من أجلك، وإن لتعاونك في تقديم المعطيات الدقيقة له الدور الأساسي في تحقيق هذا الهدف. إن المعطيات التي ستدلي بها ستكون سرية ويقتصر استخدامها على الأغراض العلمية فقط، ومن دون الحاجة لذكر اسمك.

وشكراً لتعاونك

ملاحظة: يرجى وضع علامة × في الحقل المناسب للإجابة:

1- إتجاه خط السير: من إلى

2- نوع واسطة النقل المستخدمة:

صالون (سعة 4-5 مقعد)

حافلة متوسطة الحجم (سعة 12-24 مقعدة)

حافلة كبيرة الحجم (سعة 36 مقعد فأكثر)

3- تاريخ السفر: يوم..... المصادف / / 200

4- اسم محرر الاستمارة: توقيعه

5- اسم مدقق الاستمارة: توقيعه

القسم الأول: خصائص المسافر:

1- الجنس: ذكر () أنثى ()

2- سنة الولادة:

3- معدل دخل الأسرة الشهري (دينار)

4- عدد أفراد الأسرة ()

5- المهنة: موظف () عمل حر () مزارع () غيرها () ، تذكر

القسم الثاني: هدف الرحلة

1- الذهاب أو العودة من العمل ()

2- أعمال شخصية (مراجعة دائرة مثلاً)

3- مهام وظيفية

4- أعمال تجارية أو مقاولات

5- أغراض دراسية أو تعليمية

6- زيارة الأهل أو الأقرباء

7- سياحة أو اصطيف

8- غيرها، تذكر أن يمكن

القسم الثالث: خصائص واسطة النقل

ملاحظة: يرجى تأشير مستوى الأهمية 1 أو 2 أو 3 حسب درجة القناعة لكل من

الخصائص التالية:-

مهم جداً مهم غير مهم

() () ()

1- طول الرحلة (السرعة).....

2- فترة الانتظار الواسطة.....

3- اجور النقل بالواسطة.....

4- راحة وملاءمة الواسطة

5- توفر خدمات الحجر المسبق

- 6- دقة مواعيد تحريك الواسطة
- 7- درجة أمان الواسطة من الحوادث
- 8- توفر مجال للحقائب والعفش
- 9- غيرها، تذكر إن أمكن

8-1 طرق جمع البيانات Methods Of Data Collection

بسبب تعدد طبيعة المجتمعات الإحصائية واختلاف المعطيات التي نود جمعها وظروف الإمكانيات المالية المتاحة للدراسة، فقد تعددت طرق جمع المعطيات تبعاً لذلك. وبصورة عامة هناك خمس طرق رئيسية، نستعرض فيما يأتي المفهوم العام ومجال تطبيق وخصائص كل منها:

1- طريقة الملاحظة Observation Method

وهي الطريقة التي يكون جمع المعطيات بواسطتها ممثلاً في أو معتمداً على أسلوب مراقبة الظواهر كما هي على الطبيعة، وتستخدم في حالتين:

أولاً: مراقبة الظواهر مع استخدام المنطق في تفسير ما يقع. وتستخدم عادة في بعض الحقول العلمية في دراسات اجتماعية أو تربوية أو نفسية. ومن الأمثلة على ذلك معايشة الباحث بعض فئات المجتمع لمراقبة نمط حياة هذه الفئات وما يحصل لأعضائها خلال تعاملهم ومناقشاتهم، وكما الحال عند دراسة مجتمع السجناء أو مجتمع البادية وما شابه .

ثانياً: مراقبة الظواهر لغرض التدوين (التسجيل) فقط. وفيها يقوم الباحث بمراقبة الظاهرة وتدوين الحقائق كما هي، وكما يحصل، عند وقوف الباحث مثلاً عند نقطة معينة لتسجيل حركة المرور ونمطها وذلك بتدوين عدد وسائل النقل المارة وأنواعها واتجاهها بعد تركها نقطة معينة.

مميزات وعيوب طريقة المشاهدة:

كما يتضح فإن القائمين باستخدام الحالة الأولى من هذه الطريقة هم من الكوادر المؤهلة أو المدربة جيدا والتي لها خبرة في مجال عملها، لذلك فمن المتوقع أن تقل الأخطاء مع استخدامهما، ولاسيما تلك الأخطاء التي تنتج عن غموض الهدف أو عدم وضوح مفاهيم المعطيات، بالإضافة إلى اختفاء أخطاء عدم الاستجابة. أما عيوب هذه الطريقة فتتصدر بكلفتها المرتفعة وحاجتها لكوادر مؤهلة خاصة مع الحالة الأولى.

2- طريقة التسجيل الذاتي Self-Recording Method

وتعني قيام الأشخاص المبحوثين بتدوين إجاباتهم عن الأسئلة الواردة في الاستمارة بأنفسهم. وتعد طريقة التسجيل الذاتي فاعلة في الحالات التي يكون فيها موضوع المسح والأسئلة الواردة في الاستمارة تهم المبحوثين مباشرة، كالاستفسار عن طبيعة السكن الذي يرغبون فيه أو لغرض شمولهم بإعفاءات ضريبية أو تقديم خدمات مجانية أو مخفضة لهم وما شابه، وتأخذ الطريقة عند تنفيذها واحداً أو أكثر من الأساليب الآتية:

أولاً: يقوم الباحثون بزيارة وحدات المجتمع المشمول وشرح هدف المسح الإحصائي وأهميته، ثم يتركون الاستبيانات لديهم ليقوم الأشخاص المبحوثون بملئها في وقت لاحق، ويتم الاتفاق على موعد عودة الباحثين للقيام بجمعها بعد إتمام عملية ملئها. وتساعد هذه الطريقة في التأكد من ملء الاستبيانات بشكل صحيح ودقيق.

ثانياً: ترسل الاستبيانات بواسطة البريد إلى المبحوثين للقيام بملئها، ثم يتم جمعها في وقت لاحق من قبل الباحثين أو المعنيين بالمسح.

ثالثاً: ترسل الاستبيانات بالبريد وتقوم وحدات المجتمع المشمول والتي تكون في مثل هذه الحالة غالبا مؤسسات أو شركات أو أشخاصا بملئها وإعادتها بالبريد أيضا إلى الجهة القائمة بالمسح الإحصائي. ويصلح استخدام هذه الطريقة في المجتمعات التي تقل نسبة الأمية فيها وترتفع فيها درجة

الاعتماد على البريد واستخدامه. ويفضل أن يستخدم مع هذه الطريقة كتيب يرفق مع الاستبيان لغرض المساعدة في شرحها وتوضيح كيفية ملئها.

مميزات وعيوب طريقة التسجيل الذاتي:

أ. تمتاز بانخفاض كلفتها وخاصة عند الاعتماد على البريد في إرسالها وفي وصولها ب. تحاشي تحيز الباحثين. ج. تتيح الوقت الكافي للأشخاص المشمولين بالإجابة على الأسئلة المطلوبة. د. تظهر الفائدة الكبيرة لهذه الطريقة من خلال توفير الجهود والإمكانات المالية إذا كانت وحدات المجتمع المبحوث موزعة على مناطق جغرافية متباعدة ومتعددة.

أما عيوب الطريقة فتبرز عند وجود نسبة من المشمولين لا يهتمون بإعادة الاستمارة، إما لأنهم يترددون في إعطاء بعض المعطيات بشكل صحيح، أو لصعوبة فهم الاستمارة أو بسبب الكسل في الإجابة على الأسئلة وإعادة إرسالها، ومن الممكن أيضا أن يهمل المبحوث بعض الأسئلة ويعود الاستبيان ناقصاً مما يقلل من دقة النتائج. أما العيب الآخر فهو أن الطريقة تصبح عديمة الجدوى إذا كان هناك نسبة كبيرة من المبحوثين لاتجيد القراءة والكتابة، والخدمات البريدية غير متوفرة بشكل شامل ومضمون .

3- طريقة المقابلة الشخصية Interview Method

وهي الطريقة التي بواسطتها يتم جمع المعطيات عن طريق اتصال الباحثين شخصيا بالمبحوثين لأخذ الإجابات منهم، وتعد الطريقة ملائمة للحالات الآتية:

أولاً: إذا كان عدد وحدات المشمولين صغيراً.

ثانياً: إذا كان معظم الأشخاص المشمولين أميين.

ثالثاً: إذا كانت طبيعة الاستبيان تحتاج إلى شرح وتوضيح لا يمكن فهمه عن طريق الكتيب المرفق مع الاستبيان.

إن للباحث تأثيراً كبيراً على دقة المعطيات التي تجمع بهذه الطريقة، وذلك من خلال أسلوب تعامله مع المبحوثين أثناء مقابلته لهم، لذا فمن الضروري ان تتوفر في الشخص الذي يقوم بالمقابلة الشخصية المواصفات التالية:

- أن يكون مؤهلاً لاستيعاب أهداف المسح وتعليمات الاستمارة.
- أن يكون حسن السيرة والسلوك.
- أن يتمتع بالمرونة في الحديث والقدرة على الإقناع.
- أن يتمتع بسعة الصدر والصبر والقدرة على المجاملة.
- أن يحترم العادات والتقاليد الخاصة بالأشخاص.
- أن يحترم الأسماء والألقاب الخاصة بالأشخاص.

مميزات وعيوب طريقة المقابلة الشخصية:

من مميزات هذه الطريقة أنها تساعد الأشخاص المشمولين على الإجابة من خلال قيام الباحث بتوضيح وشرح أي استفسار أو غموض، مما يساعد على زيادة دقة المعطيات وتقليل نسبة الخطأ فيها. كما تتيح هذه الطريقة للباحث التعرف على أحوال الأشخاص المبحوثين من مشاهدته مما يسهل استعمال المعطيات الخاصة بهم أحياناً. أما عيوبها فتتمثل بحاجتها إلى أعداد كبيرة من الباحثين مما يؤدي إلى زيادة كلفة المسح. بالإضافة إلى أنها قد تؤدي إلى تحيز الباحث أو قيامه بتعديل بعض الإجابات التي يسجلها من خلال التأثير الشخصي .

4- طريقة الهاتف Telephone Method

بالإضافة لما تقدم من طرق لجمع المعطيات، فإن هناك طرقاً أخرى لكنها أقل أهميته لأغلب المجتمعات النامية كطريقة الهاتف. كونها محددة للحالات التي ينتشر فيها الهاتف بصورة غالبية في المجتمع المشمول، على ان تكون المعطيات المستهدفة محدودة، وتتعلق باستطلاع آراء المبحوثين حول ظاهرة اجتماعية أو اقتصادية معينة.

5- طريقة التركيز على الآراء التي تطرح في المناقشات

الجماعية (أو العامة) Focus Group Discussion Method

وهي طريقة حديثة الاستخدام عمليا، وتتسم بالشفافية إلى حد ما، وفحواها إثارة الاهتمام بصورة غير مباشرة في التركيز على مناقشة ظاهرة أو موضوع ما في الأماكن العامة كالنوادي أو المقاهي أو أماكن العمل وغيرها لتدوين وجهات النظر التي تدلي بها الجماعة المعنية بالأمر بصورة عفوية مجردة من التأثيرات. إلا أنها قد تكون غير متوائمة لبعض أنواع المعطيات أو حتى قد غير مقبولة اجتماعيا أحيانا.

مثال (2-2): بالنسبة لمثال المسح الصناعي، يمكن استخدام طريقتين في جمع المعطيات الإحصائية هي: أسلوب المقابلة الشخصية (الأسلوب المباشر)، من خلال اتصال الباحث مباشرة بالوحدة الإحصائية (المنشأة)، ليقوم بتوجيه الأسئلة وتلقي الإجابة وتدوينها، ويتم ذلك في حالة المنشآت الصغيرة التي لا تمسك حسابات منتظمة، ولا يوجد لديها موظف مسؤول يمكنه القيام باستيفاء الاستبيان الإحصائي. أما الطريقة الثانية فهي التسجيل (التدوين الذاتي)، حيث يكون دور الباحثين هو توزيع الاستمارات على المنشآت الصناعية وفق الإطار المقرر وبمعيّتها التعليمات والتعاريف، لنقوم المنشأة بتدوين المعطيات المطلوبة، ليعود الباحث بعد ذلك ووفق موعد محدد مسبقا لاستلام الاستبيان ومراجعته بدقة عند الاستلام.

العوامل المؤثرة في اختيار طريقة جمع المعطيات:

Factors Effecting Choice of Data Collection Method

أولا: طبيعة الموضوع المراد جمع المعطيات عنه: فبينما هناك مواضيع يمكن معها اعتماد طريقة واحدة ومحددة، نجد أخرى تتطلب استخدام أكثر من طريقة. فلو افترضنا بان موضوع الدراسة يتعلق مثلا بحركة المرور أو المترددين على الأسواق العامة أو إجراء دراسة عن السجناء أو الأسعار، فمن الواضح أن طريقة المشاهدة هي الطريقة المناسبة، في حين لو كان موضوع الدراسة يتعلق مثلا بأعضاء هيئة التدريس في الجامعات أو

الموظفين العاملين في الدوائر الحكومية، فستكون طريقة التسجيل الذاتي مناسبة لذلك. أما إذا كانت الدراسة تتعلق ببعض القضايا الاجتماعية أو تخص المزارعين، وتتطلب شرح بعض الأسئلة والمفاهيم، وأن هناك نسبة من المبحوثين لا يجيدون القراءة والكتابة، فمن الأفضل اعتماد طريقة المقابلة الشخصية. وفي حالات عديدة يتطلب الأمر اعتماد أكثر من طريقة واحدة في الحصول على المعطيات، كأن نلجأ إلى المصادر التاريخية أو الوثائقية لتكوين إطار إحصائي أولاً، وإلى طريقة المقابلة الشخصية في مرحلة التنفيذ، أو كما في حالة المثال (2-2) أعلاه باستخدام طريقتي المقابلة الشخصية والتسجيل الذاتي .

ثانياً: الإمكانيات المالية والبشرية المتاحة للمسح: يعد هذا العامل من المحددات المهمة، فقد يستلزم الأمر الاعتماد على المقابلة الشخصية في جمع المعطيات وذلك لصعوبة مفاهيم الأسئلة وتعقيدها، ولكن الإمكانيات المالية المتوفرة للمسح قد تحول دون تحقيق ذلك، مما يضطرنا إلى اللجوء إلى طريقة التسجيل الذاتي مقابل القبول بدرجة دقة أقل وتوقع زيادة في نسبة عدم الاستجابة الكلية والجزئية .

9.1 اختيار وتدريب العاملين (في حالة البحوث الكبيرة).

Manpower & Training

من العوامل المهمة الأخرى التي تساعد في الحصول على معطيات دقيقة وتقلل من مسألة عدم الاستجابة الجزئية والكلية، هي عملية اختيار وتأهيل العاملين الذين يقومون بجمع المعطيات، وكذلك أولئك الذين يتولون الإشراف على المسح. ويمكن إجمال أهم المواضيع اللازم تناولها في هذا المجال بما يلي:

1. تحديد مؤهلات وعدد العاملين في المسح (في حالة البحوث الكبيرة)

غالباً ما يرتبط اختيار نوع العاملين ومؤهلاتهم وجنسهم وتحديد عددهم بطبيعة البحث وحجمه، ويفضل من هم على دراية وإلمام وخبرة بالظاهرة المدروسة. فبحث

يتعلق بالمرأة مثلا يستوجب توفير كوادرات نسائية، وتحدد مؤهلاتهم وفقا لطبيعة الأسئلة التي تحتويها الاستبيان من ناحية درجة الصعوبة والتخصص وما شابه. في حين لو كان موضوع البحث يتعلق بالبيئة مثلا، عندها سنحتاج إلى تخصصات ومؤهلات مختلفة تماما عن البحث السابق. وربما يتطلب بحث ما عدة مستويات وتخصصات في آن واحد كما في حالة البحوث متعددة الأغراض وهكذا.

2. التدريب (في حالة البحوث الكبيرة)

يتطلب تنفيذ أي بحث ميداني كبير إجراء تدريب نظري وعملي للعاملين فيه بكافة مستوياتهم، وذلك لأجل توحيد وفهم أساليب العمل وجمع المعطيات وفق مفاهيم ومصطلحات موحدة. وتتناول عملية التدريب شرح أهداف البحث وأهميته وواجبات كل من الباحثين والمشرفين وآلية العمل الميداني والمكتبي وشرح مفاهيم استبيان البحث وكيفية استيفائه، وكذلك إجراء التدريب العملي سواء بملء الاستبيان بمعطيات افتراضية أو فعلية بزيارة وحدات من المجتمع الإحصائي المشمول. ولأجل تحقيق ذلك يلزم الأمر وضع خطة للتدريب تتناول النقاط الرئيسية التالية:

أولاً: تعيين مواقع مراكز التدريب. ثانياً: تحديد عدد ومستوى كل من المتدربين. ثالثاً: تسمية القائمين بعملية التدريب. رابعاً: تحديد برنامج ومنهج التدريب ومدته الزمنية. خامساً: تحديد تكاليف ومستلزمات التدريب.

10.1. المسح التجريبي Pilot Survey

من المفيد جدا القيام بمسح تجريبي قبل التنفيذ الفعلي للبحث، يتم فيه تدريب العاملين واختبار الاستبيان الإحصائي كيما تأخذ صيغتها النهائية وتكون صالحة للتطبيق عمليا. وتتخلص عملية المسح التجريبي في توزيع عدد محدود من الاستبيانات على مجموعة من الأفراد تتشابه صفاتهم وخصائصهم مع المجتمع المعني أو بأخذ عينة عشوائية من المجتمع نفسه المراد بحثه، وذلك لتحقيق الأهداف الآتية:

- (1) إجراء تعديل إن تطلب الأمر في أسئلة الاستبيان وذلك من خلال التعرف على الواقع الفعلي للمجتمع المشمول. وقد يؤدي التعديل في الأسئلة إما إلى زيادتها أو حذف بعضها أو إدخال تغييرات عليها .
- (2) تدريب الباحثين الذين سيقومون بملء الاستبيانات قبل البدء الفعلي بالعمل الميداني والتعرف إلى المشاكل التي قد تعترض الباحثين عند أخذ إجابات من الأشخاص المشمولين.
- (3) معرفة الباحث الوقت الذي تستغرقه عملية ملء الاستبيان الواحد والاستعانة بذلك في تقدير الوقت اللازم الذي تحتاجه عملية المسح الإحصائي، وكذلك للاستعانة بذلك في تحديد عدد الباحثين المطلوب توظيفهم وفقا لمدة السح وإمكاناته المالية المتاحة.
- (4) الحصول على معلومات مفيدة للبحث، كتحديد حجم العينة وتقدير نسبة الاستجابة ووضع نظام الترميز في حالة اعتماد نظام الترميز المسبق pre-coded لعلاقة ذلك بتهيئة جداول التبويب، خاصة إذا كانت النية متجهة لاستخدام الحاسب الآلي. هذا بالإضافة إلى ما يوفره المسح التجريبي من معطيات لأغراض إدارة المسح، كتحديد الحاجة إلى وسائل نقل، والزمن الذي يستغرقه تنقل الباحثين وما إلى ذلك.

11.1. تعيين التوقيت الزمني للملائم لجمع البيانات

Survey Timing

عند اختيار الوقت الملائم لجمع المعطيات لابد من مراعاة المحددات التالية:

- (1) أن يكون الوقت متمشيا مع النمط الاعتيادي للمجتمع الإحصائي.
- (2) أن يكون ملائما للباحثين والمبجوثين. فلا نختار أكثر الأيام برودة أو أكثر الأيام حرارة مثلا، لأن ذلك سيؤثر سلبا في أداء الباحث وفي تجاوب المبجوث مع الباحث.

(3) أن نضمن وجود وحدات المجتمع، فإذا أريد مثلاً جمع المعطيات من الطلبة، فمن الطبيعي ألا نختار فترات العطل والمناسبات أو أيام الامتحانات للقيام بجمع المعطيات.

12-1. آلية العمل الميداني Field Work Processes (في حالة البحوث الكبيرة)

التحقق من مواقع الوحدات الإحصائية المشمولة في البحث. ويتم ذلك بتقسيم المنطقة الجغرافية للبحث إلى مناطق عمل رئيسية، وكل منطقة عمل يمكن ان تقسم إلى مناطق فرعية، ويكون لكل منها مجموعة بإدارة مراقب يتولى الكشف والخرطة المتعلقة بمنطقته سواء أكانت رئيسية أو فرعية، ويقوم بالتعرف على وحدات المعاينة على الطبيعة، وقد يتطلب ذلك إجراء تعديل أو إحلال للوحدات غير الموجودة أو التي وردت بطريق الخطأ.

13-1- تجهيز البيانات واستخراج النتائج Input-Output Phase

وفي هذه المرحلة تجري عملية ترقيم الاستبيانات وترميزها باعتماد أدلة خاصة بذلك، وإجراء مراجعة مكتبية لتدقيقها والتحقق من شمولية استيفاء كافة المعطيات بصورة دقيقة، ليتم بعد ذلك القيام بعملية الإدخال وإجراء عملية التدقيق النوعي validation وفق قواعد معينة تعتمد المنطق غالباً، فمثلاً لا يجوز أن يكون عمر الابن أكبر من عمر الأب، أو أن تكون المصروفات أكثر من الإيرادات وهكذا. وفي السنين الأخيرة أصبح هناك وسائل متطورة عديدة لمعالجة هذه المرحلة بسرعة ودقة عالية كما هو الحال بنظام Scanning in Data Processing System باستخدام الاستشعار البصري أو ما يطلق عليه بالذكي (Optical Intelligent or Character Recognition)

تمارين الفصل الاول

تمرين (1-1): هل لتصميم الاستبيان علاقة بأهداف المسح؟ اشرح ذلك.

تمرين (2-1): أورد مثالاً عن كيفية تحديد أهداف البحث.

تمرين (3-1): أ. ما المقصود بتحديد المجتمع الإحصائي؟ تكلم عن ذلك بإيجاز.

ب. ما فائدة تعريف وحدة المجتمع عند تصميم البحث؟

تمرين (4-1): تكلم عن مفهوم الإطار الإحصائي، وعدد الشروط اللازم توافرها فيه، مع ذكر أهم استخداماته .

تمرين (5-1): اذكر الطريقة المناسبة في جمع المعطيات لإجراء بحث للظواهر الآتية مع ذكر الأسباب:

- لدراسة عدد وجنس ووقت دخول الأشخاص أحد الأسواق العامة .
- لدراسة شمول الأطباء بخدمات اجتماعية وإعفاءات ضريبية.
- لاستطلاع آراء أعضاء الهيئة التدريسية في الكليات بشأن تطوير العملية التعليمية
- لدراسة حالة الأميين الاجتماعية والاقتصادية .

تمرين (6-1): اذكر مع الشرح المعزز بأمثلة العوامل المؤثرة على اختيار الطريقة المناسبة لجمع المعطيات .

تمرين (7-1): اشرح أهم المواصفات اللازم توافرها في الباحث عند استخدام طريقة المقابلة الشخصية لجمع المعطيات، مع ذكر أمثلة كلما أمكن ذلك.

تمرين (8-1): بين ميزات وعيوب كل من الطرق الآتية في جمع المعطيات:

- أ. طريقة المشاهدة
- ب. طريقة التسجيل الذاتي
- ج. طريقة المقابلة الشخصية

تمرين (9-1): تعد عملية تحديد المعطيات المطلوب جمعها بدقة من الأركان المهمة في إنجاح المسح الإحصائي. تكلم عن هذا الموضوع، وكيفية مراعاة هدف الدراسة في ذلك، معززا ذلك بمثال.

تمرين (10-1): وضح النقاط المهمة اللازم مراعاتها عند اختيار موعد لجمع البيانات.

تمرين (11-1): صمم نموذجاً لاستبيان إحصائي يتم تفرغ يدويا، وآخر باستخدام الحاسب، تستهدف جمع معطيات تتعلق باستطلاع آراء عينة من طلبة الجامعة عن أهمية الهوايات التي يمارسونها في أوقات الفراغ وعلاقتها بخصائص الشخص المبحوث، وذلك بهدف تحديد العوامل المؤثرة في اختيار الشباب لهواياتهم. وطبيعة المعطيات المطلوبة هي:

1- خصائص الطالب: الاختصاص والمرحلة الدراسية، العمر، الجنس، دخل الأسرة الشهري.

2- الهواية: رياضية وتشمل: قدم، سلة، منضدة، تنس، سباحة، أتنال، غيرها. غير رياضية وتشمل: طوابع، رسم، نحت، تصوير، مطالعة، سفر، زراعة ونباتات، غيرها.

3- رأي الطالب بأهمية الهواية التي يمارسها: مهمة جداً، مهمة، غير مهمة.

4- رأي الطالب في سبب ممارسة الهواية: كونها معروفة ومرغوبة في المجتمع، مرغوبة من قبل الأسرة، توفر مستلزمات ممارستها، مفيدة للصحة، مفيدة ذهنياً، رغبة شخصية، غيرها.

تمرين (12-1): صمم نموذجاً لاستبانة يتم تفرغها يدويا لدراسة كل من الظواهر الآتية:

أ. ظاهرة غياب الطلبة

ب. ظاهرة التدخين بين الطلبة

ج. ظاهرة تأخر بعض موظفي الدولة عن الدوام.

- تمرين (13-1): أ. حدد مفهوم الاستبانة والقواعد العامة لتصميمها
ب. بين شروط صياغة الأسئلة التي تتضمنها الاستبانة .
- تمرين (14-1): وضح الحالة التي تكون فيها الاستثمار مكونة من جزأين، والحالة الأخرى التي تكون فيها متكونة من ثلاثة أجزاء .
- تمرين (15-1): أ. اشرح المقصود بالمفاهيم والتصانيف الإحصائية، وأهمية اعتمادها في تحديد مفاهيم المعطيات الإحصائية
ب. تكلم بإجمال على التصنيف القياسي الدولي للأنشطة الاقتصادية.
- تمرين (16-1): أ. هناك علاقة بين نوعية الكادر المطلوب للمسح الإحصائي وطبيعة المجتمع وحجمه. تكلم عن ذلك معززا الموضوع بمثال.
ب. عدد أهم المفردات اللازم تناولها في خطة التدريب.
- تمرين (17-1): للمسح التجريبي فوائد كبيرة تنعكس على نوعية نتائج البحث. تكلم عن ذلك.



تصميم العينة

SAMPLE DESIGN

1-2- مقدمة

قبل مناقشة موضوع تصميم العينة من المفيد التطرق الى مصادر المعلومات الإحصائية المتاحة بين أيدي المخططين والباحثين والإداريين، فهي تعود إما لمصادر وثائقية (تاريخية) وتكون عادة متوفرة في السجلات والوثائق والميزانيات المالية وغيرها، والتي تتأتى من حصيلة النشاط اليومي للشركات والمؤسسات في مختلف المجالات الاقتصادية والاجتماعية والثقافية والمالية والصحية والتربوية وغيرها. فعندما تقوم هذه الشركات والمؤسسات بتهيئة هذه المعطيات وطبعتها ونشرها تسمى "بالمصادر الأولية للمعطيات"، أما عندما تقوم بتجهيز جزء من هذه المعطيات أو جميعها قبل نشرها لجهات أخرى كالمنظمات والمؤسسات الدولية أو مكاتب الإحصاء المركزية مثلا، لتتولى هذه المنظمات والمؤسسات والمكاتب طبعتها ونشرها، ففي هذه الحالة يطلق عليها "المصادر الثانوية للمعطيات". أما المصدر الآخر للمعطيات فهو يخص تلك التي يتم جمعها من مفردات المجتمع الإحصائي ميدانيا، ويكون ذلك إما من خلال شمول كافة مفردات المجتمع الإحصائي عندها يسمى "المسح الشامل أو التعداد Census" أو بشمول جزء من المجتمع الإحصائي ويطلق عليه "المسح بالعينة Sample survey". ومفهوم كل منهما هو:

1. المسوحات الشاملة (التعدادات Censuses)

وهي المسوحات التي تشمل كافة مفردات مجتمع البحث. ويقصد بالمجتمع، مجموع وحدات البحث أو الدراسة التي يراد الحصول على معطيات عنها سواء أكانت وحدة العد إنسانا أو نباتا أو جمادا. إلا أن عملية العد الكامل (المسوحات الشاملة) باهضة التكاليف سواء من الناحية المالية أو الوقت، وتتعرض لأخطاء كبيرة، كأخطاء الحذف والازدواجية، وأخطاء التسجيل، وأخطاء تجهيز المعطيات وغير ذلك التي من شأنها أن تؤثر على جودة العمل. وتتبع هذه الأخطاء أساسا من صعوبة الإشراف السليم على مثل هذه العمليات الإحصائية الكبيرة. لذا فكتيرا ما تسفر مسوحات العينة عن نتائج أكثر دقة من التعدادات، لان المسح بالعينة يتيح الوقت للإشراف الدقيق على الأعمال الميدانية وتجهيز النتائج وتقليل الأخطاء البشرية.

وشهدت السنين الأخيرة تناقصا تدريجيا في عدد المسوحات الشاملة نتيجة العوامل التالية :-

1. التطور الكبير في العمل الإداري وما أدى ذلك من انتظام السجلات الإدارية وسهولة الحصول على المعطيات الإحصائية .
2. زيادة الوعي الثقافي والاجتماعي للأفراد وإدراكهم أهمية إعطاء المعطيات الصحيحة، لازدياد حاجتهم إلى الخدمات الرسمية التي تتوقف على الحصول على معطيات مدونة عن الأفراد وممتلكاتهم وأسرهم وعناوينهم وما إلى ذلك.
3. تطور الأساليب العلمية الإحصائية والرياضية في مجال تعميم استنتاجات العينة وبناء التقديرات والتوقعات الدقيقة. وقد ساعد على ذلك بدرجة كبيرة التوسع في استخدام الحاسب الآلي.

2 المسح بالعينة Sampling Survey

إن المسح بالعينة يعني شمول جزء من المجتمع الإحصائي، على أن يكون هذا الجزء ممثلا دقيقا لخصائص المجتمع المسحوب منه هذا الجزء. ومن الأمثلة على هذا الأسلوب مسوحات تجارية وصناعية ومسح ميزانية الأسرة ومسوحات الخصوبة والظواهر الحياتية واستطلاعات الآراء حول ظاهرة معينة، قد تخص الطلبة أو إنتاج معين أو عن مستوى أو الجودة أو خدمات النقل وغيرها. ويمكن القول إن التطبيقات الرئيسية لطرق علم الإحصاء وتطوره تتم في الغالب لأغراض مسوحات العينة، وذلك لما يتمتع به هذا الأسلوب من ميزات نذكر أهمها فيما يلي:

1- توفير الوقت والجهد والتكاليف :

وتتمثل عملية توفير هذه باقتصار العمل على جزء صغير نسبيا من المجتمع الإحصائي، وهو ما يعني الحاجة إلى وقت أقل في الإعداد والتحضير للمسح ولعدد محدود من الفنيين الذين يعملون فيه، بالإضافة إلى توقع استخراج نتائج المسح في وقت أقصر بكثير مما يستغرقه المسح الشامل. إن من شأن هذا الاختصار في الجهد والوقت أن يؤدي إلى الاقتصاد في النفقات المالية للحصول على المعطيات المستهدفة.

2- زيادة دقة المعطيات الإحصائية :

قد يبدو للوهلة الأولى أن الاستنتاجات التي يتم التوصل إليها عن المجتمع من خلال دراسة نتائج العينة، هي غير مطابقة لواقع المجتمع. إلا أن استخدام الأساليب الإحصائية العلمية من قبل ذوي الخبرة والاختصاص في تصميم العينة وتقليل الأخطاء البشرية، نتيجة اقتصار الحاجة إلى عدد قليل نسبياً من الكوادر الفنية في تنفيذ مسوحات العينة، من شأنه أن يقلل كثيراً من احتمال وقوع الأخطاء وعدم قبول النتائج. بل على العكس فإن الحاجة لاستخدام أعداد كبيرة من العاملين في المسوحات الشاملة من شأنه أن يؤدي إلى تراكم أخطاء الأفراد نتيجة لتباين كفاءاتهم ومستوى تدريبهم وصعوبة متابعتهم. كما إن توافر الطرق العلمية المناسبة كقياس فترة الثقة Confidence interval واختبار الفروض Hypotheses testing وغيرها سيجب لنا فرصة التأكد من مستوى دقة النتائج وجعلنا في مأمن من معطيات العينة واستخدام نتائجها كتقديرات جيدة لمعالم المجتمع .

3- التعامل مع حالات استحالة الشمول التام

إضافة إلى ما أسلفنا فإن هناك حالات لا بد فيها من استخدام العينات حصراً، إذ لا يمكن مع تلك الحالات شمول جميع مفردات المجتمع، وذلك لما ينتج لمثل هذا الشمول من خسائر كبيرة أو بسبب الاستحالة، فمثلاً عند تحليل دم المريض يكتفي الطبيب بفحص عينة منه لأن من غير الممكن أخذ جميع دمه للاختبار، كذلك عند فحص جودة الإنتاج لا يمكن مثلاً اختبار مدى قوة مقاومة الإطار الداخلي للسيارات بتجريب كافة الوحدات المنتجة من هذه الإطارات، أو لإخضاع علب المواد الغذائية للاختبار فتح جميع العلب، لأن من شأن ذلك التسبب في خسائر مادية كبيرة وغير مبررة. كما إن هناك حالات تلزمننا باللجوء إلى العينة لاستحالة المسح الشامل معه، كما هو الحال مع المجتمعات اللانهائية مثل الطيور والأسماك وغيرها.

2-2- إجراءات تصميم العينة Sampling Design Proceses

عندما يتقرر إجراء المسح الإحصائي بأسلوب العينة، فإن ذلك يعني ان توفير المعطيات عن خصائص المجتمع سيعتمد على جزء من هذا المجتمع، ويشترط في العينة أن تكون ممثلة لخصائص مجتمع الظاهرة التي نقوم بدراستها بما في ذلك الاختلاف بين وحداته، وبحدود ما يسمح به حجم العينة تبعاً لمقياس الدقة والإمكانات المتاحة للدراسة. وفي هذا الفصل نتناول أهم الإجراءات المطلوبة لتصميم عينة والتي يتم إنجازها من خلال :

- تحديد حجم العينة Sample Size، ويراعى في اختيار أداة تحديد حجم المجتمع، وطبيعة الخاصية تحت الدراسة إن كانت على شكل نسبة أو قيمة مطلقة، وفيما إذا كان تباين المجتمع متوفراً أم لا .
- تحديد نوع العينة، بالاعتماد على طبيعة المجتمع الإحصائي وخصائصه من ناحية درجة تجانس وحداته وعما إذا كان الإطار الإحصائي للمجتمع متوفراً أم لا .
- تحديد طريقة اختيار وحدات العينة Sampling Method، والذي يعتمد على نوع العينة المقرر اختيارها، ومن بين أساليب عملية الاختيار الأسلوب الدوري periodic من خلال توظيف العينة العشوائية النظامية خاصة في حالات العينات الطبقيّة والعنقودية، هذا إضافة إلى طريقة السحب العشوائي المباشر .
- وهناك نوعان من العينات هما العينات العشوائية (الاحتمالية) والعينات غير العشوائية (غير الاحتمالية)، ولكل منها استخداماتها التي تتوقف على الغرض الرئيسي من الدراسة .

3.2 تحديد حجم العينة Sample Size

يعتبر تحديد عدد وحدات المجتمع التي ينبغي شمولها بالعينة من المسائل الأساسية في عملية تصميم العينة، وذلك لتجنب اخذ عينة صغيرة يكون تقديره للمجتمع غير دقيق وبالتالي غير مفيد .

وتتم عملية تحديد حجم العينة على مقياس تعيين درجة الدقة المستهدفة والتي يعبر عنها بحجم الخطأ المسموح به في إيجاد التقديرات والشائع يكون عند 0.05 الا انه يجب ان يعتمد على خبرة الباحث بطبيعة المجتمع المطلوب دراسته، عندها نقوم بتحديد دقة المقدر بدرجة ثقة محددة، وهي عبارة عن مقدار الاحتمال الذي يقع ضمنه تقدير معلومية المجتمع. فاذا ما افترضنا بان الخطأ المسموح به لمتوسط العينة هو 0.05 وارادنا التأكد من عدم تجاوز هذه النسبة، وعلى افتراض ان المجتمع موزع توزيعاً طبيعياً $N(0,1)$ او مقارب للتوزيع الطبيعي وبمعامل ثقة مقداره 95% فستكون لدينا فترة الثقة هي :

$$\bar{x} \pm 1.96 s/\sqrt{n}$$

حيث إن :

\bar{x} يمثل وسط العينة، وان s و n هي الانحراف المعياري للعينة و حجمها على التوالي،

s/\sqrt{n} تقدير الخطأ المعياري في المجتمع

طريقة احتساب حجم العينة

اولاً: الاحتساب عندما تكون قيم وسط المجتمع μ وانحرافه المعياري σ عبارة عن اعداد صحيحة.

1- حالة عدم معرفة حجم المجتمع الكلي N

عادة ما يراعى في اختيار أداة التحليل ان تكون كفاءة وسهولة الاستخدام. وعند مراعاة هذه الشروط، يمكن اعتماد صيغة التوزيع الطبيعي مع حالة القيم الكمية والتي تؤول الى العلاقة التالية :

نرمز الى الفرق بين x^- و μ ب d

فيكون لدينا $d = |x^- - \mu|$

ومن صيغة التوزيع الطبيعي :

$$p(|x^- - \mu| \leq Z \sigma) = 1 - \alpha$$

نحصل على :

$$d = Z \sigma$$

$$= z S/\sqrt{n}$$

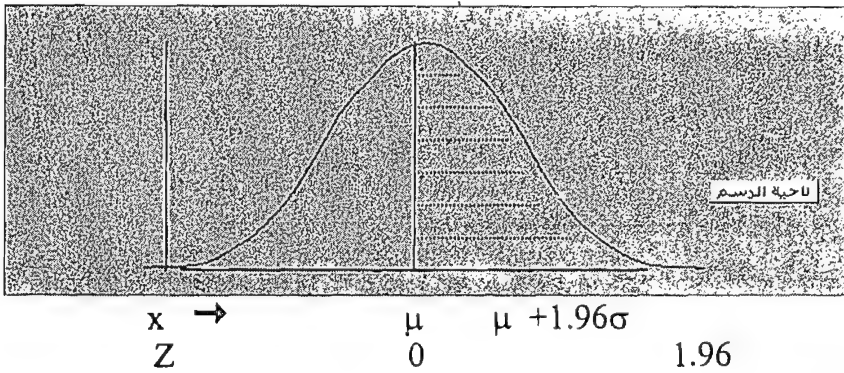
وعليه فإن :

$$n = (z s/ d)^2$$

وبما ان العلاقة تتطلب التباين S^2 فيمكن الاستعانة بإيجادها إما على نتائج مسوحات سابقة أو بإجراء مسح تجريبي بتعبئة عدد من الاستثمارات من المجتمع المشمول بالمشح. اما قيمة المتغير العشوائي الطبيعي Z فتتغير بتغير مقدار الثقة المستهدفة، فمثلا اذا كانت درجة الثقة هي 90% فهذا يعني ان درجة عدم الثقة α مقدارها $\alpha = 1 - 0.90$ ، وبالرجوع الجدول الاحصائي لتوزيع Z وعند $\alpha/2$ نجد ان قيمتها تساوي $Z(1-\alpha/2) = 1.64$. والشكل البياني (1.2) يوضح قيمة Z عند درجة ثقة مقدارها 95%.

شكل بياني رقم (1.2)

يمثل القيمة الامتسالية للمقياس الطبيعية المعيارية الواقعة بين 0 و 1.96



مثال (1.2): ما هو حجم العينة المطلوب شمولها لدراسة تقدير متوسط محصول التمخ للحيازة الزراعية الواحدة، بدرجة ثقة 95% وبفرق d بين متوسط المجتمع μ ومتوسط العينة \bar{X} لايزيد عن 1.5 كغم، ووجد من خلال مسح تجريبي ان قيمة التباين هو $S^2 = 90.3$ كغم .

الحل (1.2) :

$$n = \frac{Z^2 S^2}{d^2}$$

لدينا: $d^2 = 2.25$ ، $S^2 = 90.3$ ، $Z = 1.96$

$$\begin{aligned} n &= (3.842)(90.3)/2.25 && \text{نحصل على:} \\ &= 346.933 / 2.25 \\ &= 154 && \text{حجم العينة} \end{aligned}$$

2- حالة معلومية حجم المجتمع N فتصبح صيغة العلاقة كالآتي :

$$n = \frac{NZ^2S^2}{Nd^2 + Z^2S^2}$$

مثال (2.2): أراد احد الباحثين دراسة ظاهرة التدخين بين طلبة الجامعات، فاختار احدى الجامعات وكان عدد الطلبة فيها 6420 طالب وطالبة، واستطاع الباحث من معرفة مقدار التباين في الدخل الشهري لأسر عدد من الطلبة من خلال الاستفسار وكان مقداره $S^2 = 81$ دينار، فما هو عدد الطلبة المطلوب شمولهم في المسح، اذا كانت رغبته ان تكون درجة الثقة في المعلومات 95% ومقدار الفرق بين متوسطي المجتمع والعينة لايزيد على $d = 2$.

الحل (2.2) :

$$n = \frac{NZ^2S^2}{Nd^2 + Z^2S^2}$$

$$= \frac{(6420)(3.416)(81)}{(6420)(4)+(3.416)(81)}$$

$$= \frac{1997708.8}{259567} = 77$$

وهو حجم العينة اللازم تعبئة الاستبيانات لها .

ثانياً: حالة احتساب حجم العينة عندما يكون وسط المجتمع وانحرافه المعياري عبارة عن نسبة P (خاصية في المجتمع) :

اي في حالة ايجاد حجم العينة n لمجتمع توزيعه ثنائي Binomial distribution باحتمال P ، فإن الخطأ المعياري للتقدير P هو pq/n ، حيث ان: $q = 1 - p$ ، وان الخطأ المسموح به d ممكن ان يكون مطلقاً او نسبة. ومن الامثلة على هذا النوع من المجتمعات الاحصائية التي يعبر عنها بنسب، كنسبة المتزوجين او نسبة وحدات الانتاج الصالحة او نسبة النجاح او نسبة المواليد او نسبة قوة العمل وما شابه. ويتم الافتراض من ان توزيع هذه المجتمعات هو مقارب للتوزيع الطبيعي، وعليه يستعاض عن التباين S^2 بالمقدار pq ، فتصبح صيغة احتساب حجم العينة n على الشكل الاتي:

1- في حالة معلومية حجم المجتمع N

$$n = \frac{NZ^2(pq)}{Nd^2 + Z^2(pq)}$$

حيث إن :

P هي نسبة النجاح

q هي نسبة الفشل

مثال (3.2): يقوم مصنع لصناعة منتجات الالبان بانتاج 10000 وحدة من الجبن المتعدد الانواع يوميا، وان هناك 10% في المعدل من وحداته المنتجة هي اقل من مستوى المواصفات المحددة. فما هو حجم العينة المطلوب من خط انتاجي معين لتقدير نسبة الوحدات التي تقع تحت مستوى المواصفات المحددة، بحيث لايتجاوز الفرق في تقدير النسبة عن 0.02 وبدرجة ثقة مقدارها 90%.

الحل (3.2): عدد وحدات العينة المطلوبة

$$n = \frac{NZ^2(pq)}{Nd^2 + Z^2(pq)}$$

$$= \frac{10000(3.416)(0.10)(0.90)}{10000(0.02)^2 + (3.416)(0.10)(0.90)}$$

$$= \frac{2420.64}{4.242} = 571$$

2- في حالة عدم معلومية حجم المجتمع N

$$n = \frac{Z^2pq}{d^2}$$

مثال (4.2): قام احد اصحاب المشاتل بفحص عينة تجريبية تتكون من 48 شتلة (نبته) فوجد 15% منها مصابة بمرض، والمطلوب ايجاد حجم العينة التي يستطيع في ضوئها تحديد نسبة الشتلات المصابة في المشتل ضمن فرق مقداره 0.05 بين متوسطي المجتمع والعينة، وبدرجة ثقة مقدارها 95%.

الحل (4.2) :

$$n = \frac{Z^2pq}{d^2}$$

$$= \frac{(1.96)^2(0.15)(0.08)}{(0.05)^2}$$

أي عند درجة ثقة مقدارها 95% فإن صاحب المشتل يحتاج إلى فحص 196 شتلة لتحديد نسبة الشتلات المصابة.

ثالثاً: تحديد حجم العينة على وفق الإمكانيات المالية المتاحة
في الحالات التي تكون فيها الإمكانيات المالية محدودة ويتطلب الأمر مراعاة هذه الإمكانيات، يمكن اعتماد العلاقة التالية في تحديد حجم العينة :

$$C = c_0 - n c_1$$

$$n = (C - c_0) / c_1$$

حيث إن :

C هي الإمكانيات المالية المتاحة

C₀ نفقات الطبع والقرطاسية والتحليل وغيره من النفقات العامة

C₁ كلفة تعبئة الاستبانة الواحدة

فمثلاً إذا كان مجموع الإمكانيات المالية المتاحة هي 400 دينار وإن كلفة ملء الاستبانة الواحدة هي 5 دنانير وإن نفقات التحليل والطبع وغيرها 100 دينار، فإن حجم العينة المطلوب (عدد الاستبانات) هو :

$$n = (400 - 100) / 5 = 60$$

طبعاً في مثل هذه الحالة وكما هو الحال لو اخترنا عدد الاستبانات من دون تحديد حجم العينة مسبقاً سنحتاج إلى إيجاد "حدود الثقة" لمتوسط أحد المتغيرات الأساسية في البيانات التي يتم جمعها، وذلك بغية الاطمئنان إلى دقة حجم العينة التي تم شمولها ومن أنها ممثلة تمثيلاً صحيحاً للمجتمع المسحوبة منه ويتم ذلك باعتماد العلاقة التالية:

$$\bar{x} - z s / \sqrt{n} \leq \mu \leq \bar{x} + z s / \sqrt{n}$$

4.2 انواع العينات

اولاً: العينات العشوائية Random Samples

وهي العينات التي يتم اختيارها بطرق عشوائية وتكون مستوفية للشروط التالية :

- كل عينة يمكن اختيارها من المجتمع لها احتمال معلوم، وتبعاً لذلك فكل وحدة احتمال معلوم تشمل في العينة. وليس من الضروري ان يعني هذا الاحتمال المعلوم تساوي الاحتمال لكل وحدة في المجتمع كما هو الحال في العينات العشوائية البسيطة Simple random sample، بل قد يختلف، وهذا الاختلاف يساعد في حالة المجتمعات غير المتجانسة على توفير دقة أعلى للتقديرات التي نحصل عليها من العينة كما سيتضح عند التطرق فيما بعد إلى العينات العشوائية الطبقيّة Stratified random sample.
- تسحب العينة باستخدام إحدى طرق الاختيار العشوائي، بحيث تتحقق الاحتمالات المعلومّة.
- تعتمد الاحتمالات المعلومّة عند استخدام نتائج العينات في الحصول على تقديرات جيدة لمعالم المجتمع الذي نقوم بدراسته.

وتوجد عدة أنواع من العينات الاحتمالية، يعتمد ويتوقف استخدام كل منها على طبيعة المجتمع والغرض من الدراسة والإمكانات المتاحة، وسنتعرض فيما يلي بإيجاز إلى أهم هذه الأنواع وطرق استخدامها.

(1) العينة العشوائية البسيطة Simple Random Sample

- مفهوم العينة وشروطها: وهي العينة التي يتم اختيارها بطريقة تعطي لكل وحدة واحدة من المجتمع الإحصائي N فرصة الظهور نفسها في كل مرة من مرات الاختيار $(1/N)$ ، وبذلك فلكل عينة حجمها n احتمال الاختيار نفسه من بين العينات الممكنة أي:

$$\frac{1}{\left[\begin{matrix} N \\ n \end{matrix} \right]}$$

إذ إن الصيغة أعلاه تمثل عدد العينات الممكن اختيارها بحجم n من مجتمع حجمه N ونحصل عليها باستخدام صيغة التوافيق combination الآتية :

$$\left[\begin{matrix} N \\ n \end{matrix} \right] = \frac{N!}{n! (N-n)!} \dots\dots\dots (1.2)$$

حيث إن :

$N!$ تدعى عاملي N (مضروب N) ومفكوكه هو :

$$(N) (N-1) (N-2) \dots (2) (1)$$

مثال (5.2) : إذا كان لدينا مجتمع إحصائي متكون من الوحدات الآتية B, C, D, E :
فإن عدد العينات الممكن سحبها لحجم $n=2$ باستخدام الصيغة (1.2)
تتكون من 6 عينات هي:

BC, BD, BE, CD, CE, DE ونلاحظ أن لكل من هذه العينات نفس الاحتمال وهو $1/6$ وأن لكل وحدة في المجتمع لها الاحتمال نفسه في الظهور وهو $1/2 = 3/6$. من ذلك نستدل على أن العينة العشوائية البسيطة لها صفتان أساسيتان هما: إن لكل عنصر (أو وحده) في المجتمع احتمال الظهور نفسه، وإن لكل من العينات الست أيضا احتمال الاختيار نفسه .

- حالات استخدام العينة Sample Uses

تستخدم العينة العشوائية البسيطة عندما يكون المجتمع متجانسا من حيث الغرض أو الصفة التي تتعلق بها الدراسة، وهي أبسط أنواع العينات؛ إذ تعد أساسا لاختيار كل منها .

- أساليب اختيار العينة Sample Selection Methods

أولاً: الاختيار بالإرجاع (Selection With Replacement) ويعني أننا حين نختار مفردة من المجتمع فإننا نعيدها ثانية إلى المجتمع ليتم اختيار المفردة الثانية، وقد تظهر المفردة نفسها أو غيرها.

ثانياً: الاختيار بدون إرجاع (Selection Without Replacement) ويعني أننا عند اختيارنا للمفردة الأولى فإننا لا نلجأ إلى إعادتها ثانية إلى المجتمع وإنما نختار مفردة مما تبقى من المجتمع وهكذا. ومن الناحية العملية فإن جميع مسوحات العينة تعتمد هذا الأسلوب؛ أي بدون إرجاع .

- طريقة الاختيار العشوائي لوحدات العينة Random Selection Method Of Observations

كما هو الحال مع جميع المجالات، فقد شملت عملية التوسع في استخدام الحاسب الآلي، إجراءات السحب العشوائي لوحدات العينة، وأصبح بالإمكان في حالة إدخال معطيات المجتمع إلى الحاسوب الحصول على العينة من خلال استخدام الإيعازات التالية:

برنامج Excel ← أدوات (Tool) تحليل البيانات (Data analysis) ←
المعاينة (Sampling)

ليتم بعد ذلك الإيعاز باستخدام إحدى الطريقتين التاليتين في عملية السحب وهي إما الدورية (periodic) باعتماد (وكما أشرنا) أسلوب العينة العشوائية المنتظمة والتي تعتمد العشوائية في جزئها الأول، أو طريقة السحب العشوائي المباشر. إلا أنه في أحيان كثيرة تظهر الحاجة إلى الطريقة التقليدية في استخدام جداول الأرقام العشوائية (Random Numbers Tables) والتي تتلخص بالخطوات التالية :

أ. نعطى أرقاماً متسلسلة لجميع عناصر المجتمع ونكتب هذه الأرقام على قصاصات ورق متماثلة.

ب. نخلط هذه القصاصات خلطاً جيداً لكي نضمن ضياع أي نوع من الترتيب المحتمل بينها.

ج. تختار وحدات العينة وحدة فوحدة من بين المجموعة كلها مع الخلط الجيد في كل مرة.

د. بعد الحصول على أرقام وحدات العينة يتم تحديد وحدات المجتمع التي تحمل هذه الأرقام المختارة فنحصل على العينة المراد اختيارها من هذا المجتمع. ومن الواضح أن اتباع هذه الطريقة في كتابة الأرقام على قصاصات ورق هي غير عملية وشاقة ولا سيما إذا كان المجتمع كبير الحجم، لذا فقد أعدت جداول سميت بـ جداول الأرقام العشوائية السابق ذكرها، وتحتوي على أرقام تم الحصول عليها بطريقه عشوائية، أي بطريقه غير خاضعة لأي نوع من أنواع الترتيب، ويتم استخدامها في سحب العينات العشوائية وهي تتميز بكونها أكثر دقة وسهولة في التنفيذ من السابقة. وتتخصص طريقة استخدام جداول الأرقام العشوائية والمبين نموذج منها في الملحق رقم (1.2) بما يلي :

- نعطى أرقاماً متسلسلة لعناصر (وحدات) المجتمع المراد دراسته .
- تحديد عدد الأعمدة التي سنستخدمها من الجدول العشوائي للحصول على الأرقام المطلوبة، ويتوقف هذا على حجم المجتمع. فبذلك نختار عدد الأعمدة بحيث يكون مساوياً لعدد خانات أكبر رقم أعطي للمجتمع.
- نحدد نقطة البداية في الجداول العشوائية.
- نبدأ باختيار أول رقم من الجدول من نقطه البداية التي حددناها شرط ان يكون من ضمن الأعمدة التي اخترناها، فالعدد الذي يليه في هذه الأعمدة إلى ان نحصل على عدد وحدات العينة المطلوبة، مع استبعاد أي عدد يتكرر، أو أي عدد أكبر من عدد عناصر (مراتب Digits) المجتمع الإحصائي.
- نحدد عناصر المجتمع التي تحمل الأرقام المختارة لتكون وحدات العينة العشوائية البسيطة المراد اختيارها من هذا المجتمع .

مثال (6.2): إذا كنا بصدد القيام بدراسة عن أوضاع العاملين في أحد المصانع وكان مجموعهم 500 عامل والمطلوب اختيار عينه عشوائية حجمها 10%، حدد وحدات العينة باستخدام جداول الأرقام العشوائية.

الحل (6.2):

أ. بما أن عدد العاملين هو 500 وأن حجم العينة المطلوبة يمثل نسبة قدرها 10% فإن حجمها هو $n = 50$ عاملاً، وبذلك نعطي أرقاماً لجميع العاملين من 1 إلى 500.

ب. بما أن أكبر عدد أعطي لوحدة المجتمع هو 500 ويتكون من ثلاثة مراتب (خانات) إذن يكون عدد الأعمدة التي سنستخدمها كل مرة هو 3 أعمدة (أي أن كل عدد يتكون من ثلاثة أرقام).

ج. نحدد نقطة البداية في جدول الأرقام العشوائية، ولتكن بداية الجدول في الملحق (1.2) ولثلاث مراتب فنجد أنه الرقم 870 ولما كان هذا الرقم أكبر من 500 يتم إهماله ونأخذ الرقم الثاني وهو 48 وبما أنه أقل من 500 فإن علينا عدده الرقم الأول في العينة. ثم نأخذ الرقم الثاني المكون أيضاً من ثلاث مراتب وهو 335 وبما أنه أقل من حجم المجتمع 500 فهو يعد الرقم الثاني في العينة وهكذا حتى نحصل على 50 رقماً من بين لـ 500 دون تكرار لأي منها، وبموجب ذلك فإن أرقام العينة هي :

297، 313، 39، 65، 470، 400، 231، 250، 335، 48، 458، 340، 465،
280، 408، 405، 232، 63، 82، 297، 425، 276، 480، 350، 496، 216،
298، 233، 443، 104، 258، 382، 468، 228، 423، 397، 410، 319،
332، 287، 328، 110، 439، 487، 323، 141، 135، 191، 161، 121.

د. الآن نحدد أسماء العاملين الذين يحملون هذه الأرقام ليكونوا هم وحدات العينة العشوائية البسيطة المطلوبة.

هـ. يمكن الحصول على المعطيات المطلوبة للدراسة من هذه العينة.

و. تعمم النتائج التي نحصل عليها من هذه العينة على مجتمع العاملين بالمصنع كله وذلك باعتبار أن المعطيات التي حصلنا عليها من العينة تعد ممثلة لجميع العاملين في المصنع.

مثال (7.2): لدينا مجتمع إحصائي مكون من 50 حانوتاً (مخزناً) لبيع المواد الغذائية، وكانت قيم المبيعات اليومية (بالدينار) لهذه المخازن هي :

112، 132، 132، 131، 080، 126، 116، 118، 073، 130، 116، 120،
128، 062، 132، 091، 127، 118، 132، 132، 084، 124، 190، 109،
112، 090، 117، 127، 234، 119، 121، 128، 087، 087، 132، 129،
119، 122، 114، 093، 123، 131، 126، 112، 089، 121، 118، 116،
136، 119.

والمطلوب اختيار 10 وحدات (مخازن) كعينة عشوائية بسيطة .

الحل (7.2) :

- أ. على وفق الخطوات الواردة في أعلاه نقوم بترقيم وحدات المجتمع الإحصائي من 1 إلى 50 والتي تتكون من مرتبتين.
- ب. نستخدم الجدول في الملحق رقم (1.2) مبتدئين من السطر الأول عند العمود الثاني لتحديد وحدات العينة التي يتم سحبها. فتظهر لنا الأرقام الآتية :
- 48، 35، 49، 21، 29، 23، 44، 10، 45، 03

ج. وحسب تسلسل قيم المبيعات الواردة في المثال، نجد أن هذه الأرقام تعود إلى القيم الآتية :

116، 132، 136، 084، 234، 190، 112، 130، 089، 132

وهي تمثل وحدات العينة العشوائية البسيطة.

- عيوب العينة العشوائية البسيطة وميزاتها:

تظهر عيوب العينة العشوائية البسيطة في المجالات الآتية:

- 1- إذا كانت وحدات المجتمع غير متجانسة في الصفة التي نقوم بدراستها، فإن استخدام العينة العشوائية لا يضمن ان تكون العينة ممثلة لهذه الصفة بالمجتمع.

2- في حالة كون المجتمع الإحصائي كبيراً، فإن استخراج وحدات العينة العشوائية يحتاج إلى مجهود كبير. لتهيئة إطار المجتمع وبخاصة إذ لم نستخدم في العملية الحاسب الآلي.

3- عندما تكون وحدات العينة موزعة على مناطق جغرافية واسعة ومتباعدة، فإن تكاليف جمع المعطيات من هذه الوحدات تكون عالية عادة مع صعوبة إحكام الإشراف على العمل الميداني. وفي الواقع غالباً ما تعالج هذه العيوب باستخدام إحدى العينات العشوائية الأخرى التي سنشرحها لاحقاً.

- ميزات العينة:

كما ذكرنا فإن العينة العشوائية البسيطة تعد الأساس لباقي أنواع العينات فضلاً عن كونها من أبسط هذه العينات استخداماً.

(2) العينة العشوائية الطبقيّة Stratified Random Sample

- مفهوم العينة واستخداماتها :

لاحظنا عند التطرق إلى العينة العشوائية البسيطة أنها تستخدم مع المجتمعات المتجانسة أو قليلة الاختلاف، وبذلك نضمن الحصول على عينة ممثلة للمجتمع المسحوبة منه. أما إذا كان المجتمع غير متجانس فإن اختيار عينة عشوائية بسيطة لن يضمن ذلك. لذا نلجأ في مثل هذه الحالات إلى طريقة العينة العشوائية الطبقيّة التي نتعامل مع المجتمعات غير المتجانسة.

وتتلخص خطوات اختيار وحداتها بما يلي :

الخطوة الأولى، وفيها يقسم المجتمع غير المتجانس إلى مجتمعات صغيرة $N_1, N_2, N_3, \dots, N_k$ تكون متجانسة بالنسبة للصفة التي نقوم بدراستها، كأن تكون هذه الصفة هي العمر أو الدخل أو غيرها، على أن لا يحصل تداخل بين وحداتها، أي لا تتكرر الوحدة نفسها في أكثر من طبقة واحدة، بحيث يتحقق

$$N = N_1 + N_2 + N_3 + \dots + N_k$$

وفي الخطوة الثانية، نختار عينة عشوائية بسيطة من كل طبقة، بحيث تكون العينة المختارة من الطبقات المختلفة هي العينة العشوائية الطبقيّة اي ان:

$$n = n_1 + n_2 + n_3 + \dots + n_k$$

- طرق تحديد عدد وحدات العينة التي يتم سحبها من كل طبقة

اولا: طريقة الاختيار التناسب Oproportional allocation method

وبموجب هذه الطريقة فان حجم العينة لكل طبقة يكون متناسبا مع نسبة حجم الطبقة الى الحجم الكلي للمجتمع الاحصائي. اي ان حجم العينة العشوائية المأخوذة من طبقة ما الى حجم العينة النهائي يكون مساويا لحجم تلك الطبقة الى الحجم الكلي للمجتمع. ويمكن التعبير عن ذلك بالصيغة الآتية:

$$W_i = \frac{N_i}{N} = \frac{n_i}{n}$$

حيث إن :

W_i هي نسبة العينة i الى حجم العينة الكلي وبهذا يكون حجم العينة i من الطبقة i هو :

$$n_i = n(N_i / N)$$

حيث إن :

$$\sum n_i = n \text{ حجم العينة الكلي، اي}$$

$$\sum N_i = N \text{ حجم المجتمع الكلي، اي}$$

مثال (8.2): لنفترض ان لدينا مجتمعا يتكون من 25 أسرة وان المصروفات الشهرية الاسبوعية بالدينار لكل من هذه الاسر هو كما مبين في الآتي، والمطلوب سحب عينة عشوائية طبقية تتكون من 8 أسر باستخدام طريقة الاختيار المتناسب .

48, 43, 44, 19, 16, 14, 18, 12, 17, 15, 10, 46, 42, 38, 45, 41, 40, 50, 32, 23, 30 29, 24, 26, 24.

الحل (8.2) :

من ملاحظة ارقام المجتمع الاحصائي نستدل على امكانية تقسيم المجتمع الى ثلاث طبقات، قيمها هي :

الطبقة 1 (N1): 10, 15, 17, 12, 18, 14, 16, 19

الطبقة 2 (N2): 32, 23, 24, 26, 27, 29, 30

الطبقة 3 (N3): 50, 40, 41, 45, 38, 42, 46, 44, 43, 48

اي: $N1 = 8, N2 = 7, N3 = 10$

وباستخدام الصيغة اعلاه $n_i = n(N_i / N)$

نحصل :

$n_1 = 8 (8/25) = 2.56 \approx 3$ وهي عدد وحدات عينة الطبقة N1

$n_2 = 8 (7/25) = 2.24 \approx 2$ وهي عدد وحدات عينة الطبقة N2

$n_3 = 8 (10/25) = 3.2 \approx 3$ وهي عدد وحدات عينة الطبقة N3

وفي المرحلة الاخيرة نستخدم الجداول العشوائية على وفق الخطوات الواردة

في أعلاه. فنحصل على وحدات العينة التي ظهرت من كل طبقة على النحو الآتي:

n_1 : 14, 17, 10

n_2 : 27, 23

n_3 : 38, 41, 44

وبذلك فان وحدات العينة n هي: 38, 41, 44, 27, 23, 10, 17, 14

ثانيا: طريقة الاختيار الامثل Optimal allocation method

وتقوم هذه الطريقة على اساس تقليل التباين، وعلى افتراض ان تكاليف اختيار

الوحدة متساوية، فان صيغة العلاقة يمكن التعبير عنها كما يلي :

$$n_i = n \frac{N_i S_i}{\sum N_i S_i}$$

وتدعى هذه العلاقة ايضا بالاختيار الامثل لنيمان (Nymen)، حيث ان n هي

حجم العينة الطبقة S_i و S_i هو الانحراف المعياري .

مثال (9.2): يوجد في إحدى المزارع 34 بقرة، كمية انتاج كل منها من الحليب (كغم) هي كما مبين في الآتي، والمطلوب اختيار عينة طبقية عدد وحداتها $n = 8$ باستخدام طريقة الاختيار الامثل.

82, 81, 81, 76, 85, 88, 67, 63, 56, 57, 56, 53, 57, 61, 62, 62, 69,
60, 59, 51, 54, 53, 51, 78, 87, 98, 96, 95, 85, 89, 74, 76, 75, 62,

الحل (9.2) :

أ. نقسم المجتمع الاحصائي الى طبقتين، ونصنف الابقار التي كمية انتاجها يقل عن 50 كغم في الطبقة الاولى N_1 وتلك التي يبلغ انتاجها 70 كغم فاكثر في الطبقة الثانية N_2 ، وبذلك يصبح لدينا :

- الطبقة الاولى $N = 18$ قيم وحداتها هي: 53, 56, 62, 69, 60, 59, 67, 65, 63, 56, 57, 51, 54, 53, 51, 62, 61, 57
- الطبقة الثانية $N = 16$ وقيم وحداتها هي: 98, 96, 95, 85, 76, 88, 87, 89, 82, 81, 81, 85, 75, 75, 74, 76

ب. نستخرج الانحراف المعياري لوحداث كل من الطبقتين وكالاتي :

$$S_1 = \sqrt{[\sum X^2 - (\sum X/N)^2] / N}$$

فبالنسبة للطبقة الاولى يصبح لدينا :

$$\sum X = 1056 \quad \sum X^2 = 62440 \quad 1 = 18$$

$$S_1 = \sqrt{27.111} \\ = 5.21$$

اما الطبقة الثانية فلدينا:

$$\sum X = 1345 \quad \sum X^2 = 114112 \quad N_2 = 16$$

$$S_2 = \sqrt{55.98} \\ = 7.42$$

ج. نحدد عدد الوحدات اللازم سحبها من كل طبقة على النحو التالي :

$$n_i = n \frac{NiSi}{\sum NiSi}$$

$$n_1 = 8 \frac{(18)(5.21)}{(8)(5.21) + (16)(7.42)} = 3.53 \approx 4$$

وهي عدد الوحدات المطلوب سحبها من الطبقة الاولى N1

$$n_2 = 8 \frac{118.656}{212.31} = 4.47 \approx 4$$

وهي عدد الوحدات المطلوب سحبها من الطبقة الثانية N2

وعليه فإن مجموع وحدات العينة العشوائية الطبقيّة هو: $n = n_1 + n_2 = 4 + 4 = 8$

وبالاختيار العشوائي فان الوحدات التي قد تظهر لنا هي: 53, 57, 65, 54, 88, 95, 76, 78 .

(3) العينة العشوائية المنتظمة Systematic Random Sample

- مفهوم العينة واستخداماتها :

اولا: في حالة عدم معلومية حجم المجتمع

لقد تعرضنا الى المعاينة العشوائية البسيطة في حالة المجتمعات المتجانسة والى المعاينة العشوائية الطبقيّة في حالة المجتمعات غير المتجانسة؛ اذ يتطلب كل منهما الى معرفة حجم المجتمع وغالبا ما نكون مكلفتين من ناحية الجهد والوقت والتكاليف، اما في حالة المعاينة العشوائية المنتظمة فهي العينة المناسبة للاستخدام عندما لا نتمكن من تحديد حجم المجتمع الذي نقوم بدراسته، وتتلخص في اختيار كل i th وحدة على التوالي بعد تحديد نقطة البداية عشوائيا بين الاعداد من 1, 2, ... i وبسبب اختيار وحدات العينة بطريقة منتظمة بعد نقطة البداية جاءت تسميتها بالعينة العشوائية المنتظمة، فاذا اردنا مثلا اختيار عينة باختيار كل عاشر وحدة فان علينا ان نحدد

نقطة البداية عشوائيا من بين الاعداد 1 و 10 وليكن الرقم 4، حينئذ تكون وحدات العينة المنتظمة هي: ... 4, 14, 24, 34, ولغاية الحصول على عدد وحدات العينة المطلوبة. والعينة العشوائية المنتظمة كثيرة الاستخدام في المجالات التطبيقية، فقد يتم مثلا اختيار عينة منتظمة من انتاج آلة لمراقبة الجودة، او عينة من الترددات على مكتبة عامة او على مصرف او مستشفى، او اختيار عينة ميدانية من المساكن او المتاجر او وسائل المارة وغير ذلك. ويأتي كثرة استخدام هذا النوع من العينات لميزاتها في تقليل التكاليف وسهولة التطبيق. كما أن وحداتها تتوزع توزيعا منتظما أكثر مما يحصل مع العينة العشوائية البسيطة التي قد تتركز الوحدات فيها في موقع واحد .

ثانيا: في حالة معرفة حجم المجتمع :

عند معلومية حجم المجتمع N فان اختيار وحدات العينة العشوائية المنتظمة بحجم n يتم على النحو الاتي :

- أ. نحدد طول دورة المعاينة ولنرمز لها L اذ إن: $L = N / n$
- ب. نحدد نقطة البداية وذلك بأختيار رقم عشوائيا على ان يقع بين 1 و L
- ج. نضيف في كل مرة طول الدورة L الى الرقم الذي تم اختياره، لغاية الحصول على حجم العينة n المطلوب. فإذا اردنا مثلا اختيار عينة عشوائية منتظمة بحجم $n = 10$ من مجتمع يتكون من 100 وحدة فيتم ذلك كالآتي:

$$L = 100 / 10 = 10 \text{ طول الدورة}$$

نحدد نقطة البداية عشوائيا بين 1 و 10 ولتكن الرقم 4
نحدد وحدات العينة باضافة طول الدورة 10 الى الرقم الاول وهو 4 بانتظام
أي: $4, 4+L, 4+2L, 4+3L, \dots, 4+(n-1)L$

فنحصل على وحدات العينة التالية:

$$4, 14, 24, 34, 44, 54, 64, 74, 84, 94$$

- عيوب العينة العشوائية المنتظمة :

للعينة عيبان، احدهما حاصل والثاني محتمل الوقوع وهما :
فالعيب الحاصل يتمثل في انه لا توجد للعينة طريقة ذات اعتمادية عالية في تقدير الخطأ المعياري لمتوسط المجتمع رغم شمولها ضمنيا على طبقات، لأن العشوائية تحصل مع المفردة الاولى لكل طبقة، وهي بذلك تختلف عن العينة العشوائية الطبقيّة .

أما العيب المحتمل وقوعه فيحصل عندما تأخذ وحدات المجتمع نسقا دوريا، كما هو الحال عند الرجوع الى قوائم السكان حيث نجد ان ترتيب افراد الأسرة يبدأ برب الأسرة ومن ثم الزوجة فالاولاد الاكبر فالاصغر وهكذا، ففي مثل هذه الحالة تمثل الوحدة الاولى عند ظهورها دائما رب الأسرة، والثانية غالبا الزوجة، والثالثة غالبا الابن الاكبر وهكذا. وعليه فاذا كان ترتيب وحدات المجتمع موضع الدراسة ترتيبا دوريا فلا يصح الاستعانة بهذا النوع من العينات.

(4) العينة العشوائية العنقودية Cluster Random Sample

- مفهوم العينة واستخداماتها :

بصورة عامة يمكن القول بان انواع العينات الثلاث السابق ذكرها هي الاكثر استخداما وانتشارا على نطاق المسوحات الاحصائية الميدانية التي يقوم بها الباحثون شخصا او تلك التي تقوم بها المنظمات. إلا اننا نلاحظ في بعض الدراسات التطبيقية ان وحدات بعض المجتمعات توجد على شكل تجمعات غالبا ما تكون متشابهة الى حد كبير بالنسبة للخاصية التي نقوم بدراسة مثل: المدن، الشوارع، الكليات، المناطق الزراعية وغيرها، وتسمى هذه التجمعات بالعناقيد Cluster اذ يحتوي كل عنقود على عدد من عناصر المجتمع الاصلية التي غالبا ما تكون متجانسة. وعادة ما يستخدم مع هكذا حالات طريقة العينة العشوائية العنقودية، ويأتي استخدامها لسببين رئيسيين هما:

أ. عدم توفر اطار احصائي دقيق للمجتمع، او ان كلفة توفيره تكون باهظة التكاليف. فمثلا لو كنا بصدد اجراء مسح اقتصادي واجتماعي وان وحدة المشاهدة فيه هي الاسرة، لكن قائمة باسماء الاسر لم تكن متوفرة، بينما تتوافر قائمة باسماء الاحياء (المناطق) وهي متشابهة من ناحية الخاصية التي نقوم بدراستها (مثال مستوى الدخل - مستوى المعيشة... الخ) وكل من هذه الاحياء تضم مجموعة أسر، ففي مثل هذه الحالة يمكن اختيار عينة عشوائية من الاحياء، ومن ثم اخذ عينة من الأسر من الاحياء المختارة .

ب. لتركيز الجهود والاموال مما يساعد في الوصول الى وحدات المجتمع بكلفة وجهد اقل مما عليه في حالة العينات السابق ذكرها.

- اسلوب اختيار العينة :

إن اختيار العينة العشوائية العنقودية يتم إما على مرحلة واحدة وذلك باختيار عينة عشوائية بسيطة من العناقيد ثم دراسة وحدات هذه العناقيد، او باكثر من مرحلة واحدة، اذ نقوم مثلا باختيار عينة عشوائية من العناقيد في المرحلة الاولى، بعدها يتم اختيار عينة عشوائية بسيطة من كل عنقود مختار في المرحلة الثانية لتكون بذلك قد تمت بمرحلتين. او تكون باكثر من مرحلتين خاصة اذا كان المجتمع يتصف بالتجانس، كما لو كان لدينا مجتمع الريف مثلا فاننا نقوم في المرحلة الاولى باختيار عينة عشوائية من القرى، ومن القرى المختارة نختار عينة عشوائية من المدارس في المرحلة الثانية، ومن ثم نختار عينة عشوائية من طلبة هذه المدارس في المرحلة الثالثة وهكذا. وهو ما يطلق عليه بالعينة العشوائية متعددة المراحل Multi-Stage Random Sample.

ثانيا: العينات غير العشوائية Non-Random Samples

(1) العينة المتعمدة (او التحكمية) Judgement Sample

وهي العينة التي يتم اختيار وحداتها وفق وجهة نظر الباحث لاعتقاده من انها تعطي نتائج مرضية.

(2) العينة الحصصية Quata Sample

وبموجبها يتم ايضا اختيار وحدات العينة وفق وجهة نظر الباحث ولكن تركيبها يكون حسب نسب الاجزاء الموجودة بالمجتمع، فاذا كان المجتمع يتكون مثلا من ثلاث فئات من دخول الأسرة ولنقل: أسر ذات دخل متدنٍ، واسر متوسطة الدخل، وثالثة هي أسر عالية الدخل. وكانت نسب كل من هذه الفئات في المجتمع هي 40% ، 50% ، 10% على التوالي، فان اختيار عينة تتكون من 1000 أسرة يجب ان تضم ذات النسب المذكورة، بحيث تشتمل على 400 أسرة من ذوي الدخل المتدني و 500 أسرة من ذوي الدخل المتوسط و100 من ذوي الدخل العالي. إلا ان عملية الاختيار تتم بصورة كيفية من دون الاعتماد على الاسلوب العشوائي .

للتوسع في فهم العينات واسلوب تقدير معالم المجتمع باستخدام نتائجها وكذلك لمعرفة حساب الاخطاء المعيارية لكل منها وتقدير مجموع المجتمع، بالاضافة إلى كيفية ايجاد فترة الثقة لكل من متوسطات وتباينات هذه العينات يمكن الرجوع الى كتاب المؤلف: الطرق الاحصائية التطبيقية للمعاينة، جامعة السابع من ابريل- ليبيا، 1995.

تقارین الفصل الثانی

تمرین (1-2): اشرح الاسباب المؤدية الى ضرورة تحديد حجم العينة للمسح الاحصائي.

تمرین (2-2): ان حجم الخطأ المسموح به وحدود الثقة المقررة هي من العوامل المحددة لمستوى دقة العينة، اشرح المقصود بكل من هذين المفهومين.

تمرین (3-2): اذا كان لدينا مجتمع احصائي يتكون من 500 بقرة، وكان تباين انتاج البقرة من الحليب هو 20.1 كغم شهريا، أوجد حجم العينة المطلوبة من الابقار لدراسة اسباب الاختلاف في انتاج الحليب، بدرجة ثقة مقدارها 95%، وبفرق بين متوسطي المجتمع μ والعينة \bar{x} مقداره 2 كغم.

تمرین (4-2): معمل للصناعات الجلدية يقوم بانتاج 8000 حقيبة جلدية مدرسية خلال الشهر، ووجد من خلال عينة تجريبية ان مانسبته 5% من هذه الحقائب غير صالحة، وبغية دراسة اسباب الخلل في الانتاج، تقرر اخذ عينة عشوائية، فما هو حجم العينة المناسب عند مستوى معنوية مقداره 0.10 وبفرق مقداره 0.05 بين متوسطي المجتمع والعينة.

تمرین (5-2): من خلال فحص دفاتر الامتحان النهائي، وجد ان 20% من الطلبة الممتحنين لم يحققوا درجة النجاح، فما هو حجم العينة المطلوبة التي يتسنى في ضوئها دراسة الحالة، ضمن فرق مقداره 0.03 بين متوسطي المجتمع والعينة، وبدرجة ثقة مقدارها 95% .

تمرين (2-6): ما هو حجم العينة المناسب، إذا كانت المكانات المالية المتاحة هي 1500 دينار، وان كلفة تصميم المسح الميداني تقدر بحوالي 650 دينار، وكلفة ملء الاستبانة الواحدة واستخراج نتائجها تبلغ 9 دنانير .

تمرين (2-7): أ. وضح خطوات تصميم العينة العشوائية الطبقية.
ب. اشرح طريقة الاختيار المناسب مع ذكر صيغة احتسابها.

تمرين (2-8): في ادناه قيم وحدات مجتمع احصائي، والمطلوب اختيار عينة عشوائية منتظمة تتكون من 4 وحدات. 51, 62, 59, 76, 63, 71, 54, 61, 70, 50, 58, 57, 60, 65, 72, 68, 74, 73, 66, 55

تمرين (2-9): اتضح من احدى المؤسسات التي تضم 3600 موظف ان نسبة المتأخرين عن موعد الدوام الرسمي تصل في المعدل الى 4.6%، فما هو حجم العينة المطلوبة لدراسة اسباب هذه الظاهرة، على ان لايتجاوز الفرق في نسبتي العينة والمجتمع عن 0.01 وبدرجة ثقة مقدارها 95%.

تمرين (2-10): لاعداد دراسة عن الحالة التعليمية للاناث في مدينة ما. اختير احد الاحياء الذي يضم 60 أسرة فكانت نسبة الاناث 0.55 وكان توزيع المجتمع مقارب للتوزيع الطبيعي، فما هو حجم العينة اللازم سحبها بفرق 0.01 بين متوسطي المجتمع والعينة وبدرجة ثقة مقدارها 90%.



تبويب و عرض البيانات

DATA TABULATION & PRESENTATION

3.1- مقدمة

عقب مرحلة جمع البيانات والمعلومات الاحصائية ميدانيا بواسطة الاستبانات او عند نقل معلومات من السجلات والوثائق، يصبح من المطلوب تهيئتها على شكل جداول بالصيغة التي تمكننا من الاطلاع على اتجاهها وعلى مدلولاتها، وبما يساعد على استخدامها لاجراض التحليل للكشف عن طبيعة العلاقة بين متغيراتها. ولهذا الغرض فالخطوة الاولى المطلوبة هي وضع بيانات كل استبانة أو مشاهدة (observation) أو مجموعة مشاهدات (فئة) في صف (سطر) واحد، ويشمل ذلك القيام بتحويل البيانات النوعية (غير الرقمية) الى بيانات كمية (رقمية) أو اعادة صياغتها بالشكل الذي يفي بحاجة عملية التحليل. في الآتي نتناول الاجراءات المطلوب اتخاذها بهذا الاتجاه مبتدئين باستخدام برنامجي SPSS بصورة اساسية وبرنامج EXCEL لحالات اضافية محدودة، ومن ثم العروج على كيفية القيام بانجاز ذلك يدويا من دون استخدام الحاسوب لمعرفة اسس نتائج الحاسوب.

وسيعتمد استخدام برنامج EXCEL في حالات محدودة يكون فيها اكثر سهولة وكفاءة مع الحالة التحليلية المطلوبة كما هو الحال مع العرض البياني، وبصورة عامة يمكن القول بان برنامج SPSS هو اكثر ملاءمة في الحالات التحليلية التي نحتاج فيها الى تفصيل وعمق اكثر في المخرجات لاثبات مدى معنوية النتائج كما هو الحال في موضوع تحليل الانحدار Regression Analysis وموضوع تحليل التباين Variance Analysis ومصفوفة الارتباط Correlation Matrix وفي ادخال وتبويب البيانات لغرض التدقيق والحصول على مقاييس النزعة المركزية والتشتت وشكل توزيع البيانات (الالتواء skewness والتفرطح kurtosis) واختبار الفرضيات Hypotheses Testing، في حين ان مخرجات برنامج Excel المجملة هي وافية وكفوة لمواضيع التوزيع التكراري على شكل فئات، وفي حالة الرسوم والعرض البياني التي يتوفر لها في البرنامج العديد من خيارات ومزايا العرض بالاضافة الى خاصية تسهيلات العمليات الحسابية التي قد يحتاجها الباحث من خلال الجداول الالكترونية وشريط الصيغ.

3.2- ادخال البيانات باستخدام برنامج SPSS

وحيث يتعذر اخضاع المتغيرات الفئوية (غير الرقمية) للتحليل العلمي فمن المفيد الإشارة الى انه بالامكان القيام بتحويل البيانات غير الرقمية إلى قيم رقمية (كمية) قبل الادخال أو لاحقاً بعد الادخال باستخدام الامر Transferring الذي سنتطرق اليه لاحقاً، ويتم ذلك باعطاء رمز رقمي بدلاً من الاجابات غير الرقمية، فمثلاً اذا كانت الاجابة على احد الاسئلة: موافق جداً - موافق - غير موافق، تصبح لاغراض التحليل (3 - 2 - 1) اي تعطى القيمة (3) للجابة بموافق جداً والقيمة (2) للجابة لموافق والقيمة (1) للجابة غير موافق وهكذا. وعادة ما يطلق على متغيرات هذا النوع من القيم الجديدة بالمتغيرات الهيكلية (Dummy Variables).

وتجرى عملية الادخال بشكل متسلسل، فكل سطر أو صف تعود بياناته لمشاهدة معينة (استبانة أو شخص)، وكل موقع في السطر يخص متغيراً محدداً وهكذا. وفي حالة مصادفة وجود بيانات مفقودة لمتغير أو أكثر يترك مكانها خالياً ليتم معالجتها بعد الانتهاء من عملية الادخال، من خلال القيام بإجراء التقدير أو التعويض لكي يبقى كل عمود خاص بمتغير محدد وكل موقع في العمود يعود لمشاهدة محددة، على ان يحمل كل متغير اسماً أو رمزاً، وفي الغالب ما يرمز للمتغيرات بـ $X_1, X_2, X_3, \dots, X_k$.

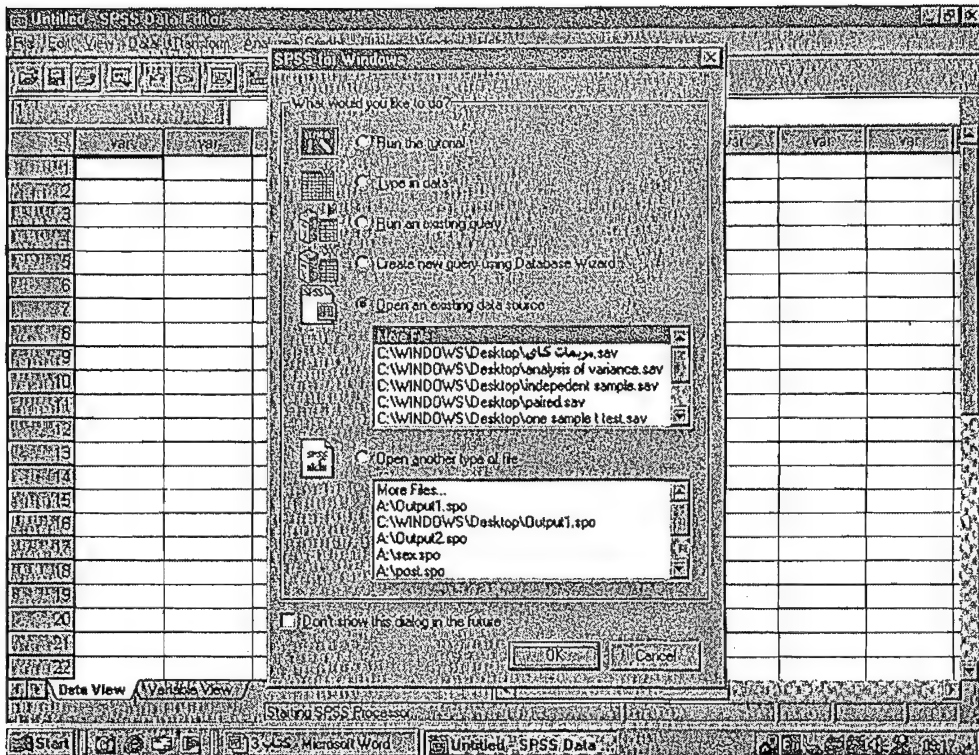
مثال (1.3): لنفترض لدينا استبانات تم جمعها من عينة شملت 31 طالباً لدراسة مدى تأثير عوامل محددة على مستوى اداء الطالب في امتحان الاحصاء، وكانت الاسئلة التي تضمنتها الاستبانة هي: الدرجة النهائية في امتحان الاحصاء، معدل الثانوية العامة، الفرع الدراسي في الثانوية العامة، الجنس، العمر، التحصيل العلمي للاب.

فلتوبيب الاجابات الواردة في الاستبانات باستخدام برنامج SPSS نقوم أولاً بالدخول الى البرنامج وفق التسلسل التالي :

الامر الفرعي (New) → الامر الرئيسي (File) → spss → program → start

والطريقة الثانية عند الدخول الى البرنامج تظهر لنا لوحة تحمل قائمة بالخيارات ان كان الامر هو استخدام احد الملفات المتوفرة، أو ادخال بيانات لتكوين ملف جديد، كما هو مبين في الشكل رقم (3-1) ادناه :

شكل رقم (3-1) بوضع الخيارات التامة عند الدخول الى برنامج SPSS



ف نقوم بالتأشير على موقع Type in data ومن ثم الكيس على ايقونة Ok الموجودة في اسفل القائمة فتظهر الصفحة التي يتم فيها تدوين اسماء المتغيرات المزمع تبويب بياناتها والمبينة في الشكل (3-2) ادناه والتي تحمل عنوان variable view المدونة في اسفل الجدول :

الشكل رقم (2.3) جدول Variable View

| Name | Type | Width | Decimals | Label | Value | Missing | Columns | Categorical |
|------|---------|-------|----------|-----------------------|-------|---------|---------|-------------|
| Y | Numeric | 8 | 0 | درجة الإحصاء النهائية | None | None | 8 | Cer |
| x01 | Numeric | 8 | 0 | المتغير | None | None | 8 | Cer |
| x02 | Numeric | 8 | 0 | المتغير | None | None | 8 | Cer |
| x03 | Numeric | 8 | 0 | معدل الثانوية العامة | None | None | 8 | Cer |
| x04 | Numeric | 8 | 0 | معدل الثانوية العامة | None | None | 8 | Cer |
| x05 | Numeric | 8 | 0 | شهادة الإحصاء | None | None | 8 | Cer |
| x06 | | | | | | | | |
| x07 | | | | | | | | |
| x08 | | | | | | | | |
| x09 | | | | | | | | |
| x10 | | | | | | | | |
| x11 | | | | | | | | |
| x12 | | | | | | | | |
| x13 | | | | | | | | |
| x14 | | | | | | | | |
| x15 | | | | | | | | |
| x16 | | | | | | | | |
| x17 | | | | | | | | |
| x18 | | | | | | | | |
| x19 | | | | | | | | |
| x20 | | | | | | | | |
| x21 | | | | | | | | |
| x22 | | | | | | | | |
| x23 | | | | | | | | |

وكما مبين من الشكل (2-3) اعلاه فقد تم تدوين اسماء المتغيرات على شكل رموز تحت عمود Name في حين تم تدوين المقصود بكل من هذه الاسماء أو الرموز في عمود Label فمثلا اعطاء الرمز Y في عمود Name وتعريف Y في عمود Label من انه "علامة الاحصاء النهائية" واعطاء الاسم X01 بعد ذلك في عمود Name وتعريفه في عمود Label على انه متغير "معدل الثانوية العامة" وهكذا. اما الحقول الاخرى من الجدول فيمكن استخدامها وفق الحاجة كتغيير عدد المراتب الكسرية أو ما يتعلق بعرض العمود أو تحديد نوع المتغير اهو كمي أو نوعي...الخ. وبعد الانتهاء من تدوين اسماء كافة المتغيرات المطلوب تبويبها في الجدول، نقوم بالكبس (click) على Data View المبينة في اسفل الصفحة ايضا ليظهر لنا جدول وهو يحمل اسماء المتغيرات التي تم تدوينها في Variable View. وفي الجدول Data View والمبين في الشكل (3.3) التالي يتم فيه ادخال بيانات الاستبانة البالغ عددها 31 استبانة وتضم 6 متغيرات، وللتمكن من اجراء عمليات التحليل الاحصائي على البيانات يستوجب وكما ذكرنا تحويل المتغيرات النوعية الى كمية، ولهذا الغرض

فبالنسبة لمتغير "الجنس" فقد تم اعطاء القيمة 1 للذكور والقيمة 2 للاناث، كما وتم اعطاء القيمة 1 للاختصاص الادبي و 2 للاختصاص العلمي في متغير "الاختصاص في مرحلة الثانوية العامة"، اما فيما يخص متغير "شهادة الاب" فقد اعطيت القيم 1، 2، 3، 4، 5 لشهادة الابتدائية، الثانوية، الاعدادية، الجامعية، شهادة عالية على التوالي.

الشكل (3.3): جدول Data View

| | X01 | X02 | X03 | X04 | X05 | X06 |
|----|-----|-----|-----|-----|-----|-----|
| 1 | 41 | 2 | 20 | 61 | 1 | 3 |
| 2 | 40 | 2 | 22 | 70 | 2 | 4 |
| 3 | 91 | 2 | 21 | 71 | 2 | 4 |
| 4 | 75 | 2 | 23 | 69 | 1 | 3 |
| 5 | 75 | 1 | 20 | 65 | 2 | 3 |
| 6 | 64 | 1 | 22 | 59 | 1 | 3 |
| 7 | 59 | 1 | 25 | 59 | 1 | 3 |
| 8 | 42 | 2 | 2 | 56 | 1 | 3 |
| 9 | 56 | 1 | 23 | 60 | 1 | 3 |
| 10 | 52 | 2 | 24 | 65 | 2 | 3 |
| 11 | 50 | 2 | 20 | 68 | 1 | 3 |
| 12 | 95 | 1 | 21 | 76 | 2 | 5 |
| 13 | 61 | 1 | 23 | 72 | 2 | 4 |
| 14 | 68 | 2 | 53 | 65 | 1 | 4 |
| 15 | 63 | 1 | 20 | 59 | 2 | 4 |
| 16 | 65 | 2 | 25 | 82 | 2 | 1 |
| 17 | 68 | 2 | 22 | 60 | 1 | 2 |
| 18 | 70 | 1 | 23 | 72 | 1 | 5 |
| 19 | 60 | 1 | 22 | 70 | 2 | 4 |
| 20 | 63 | 1 | 21 | 60 | 2 | 4 |
| 21 | 64 | 1 | 20 | 81 | 2 | 4 |
| 22 | 68 | 2 | 20 | 83 | 2 | 3 |

وفي حالة عدم الاجراء اللاحق لعملية تحويل المتغيرات من قيم غير رقمية (نوعية) الى قيم رقمية (كمية) وكذلك في حالة الحاجة لاعادة صياغة بعض المتغيرات، يمكن اللجوء الى الامر الرئيسي Transfer من برنامج SPSS ومن بين الاوامر الفرعية التي يمكن اللجوء اليها في هذا الامر مثلا هي :

- الامر الفرعي Recode: لاعادة ترميز المتغير المعني.
- الامر الفرعي Compute: ويستخدم لتشكيل متغيرات جديدة اعتمادا على قيم المتغيرات المتوفرة، والتي يمكن ان تتطلب اجراء عمليات حسابية يقوم بانجازها، كما يمكن الاستعانة بهذا الامر الفرعي لايجاد قيم تقديرية للقيم المفقودة. فمثلا لتشكيل متغير جديد وليكن معدل دخل الفرد في الأسرة ولدينا

مجموع دخل الأسرة وعدد افرادها فنحصل على المتغير الجديد بقسمة مجموع الدخل على عدد الافراد.

ومن الاجراءات التي تتضمنها الاوامر الفرعية اعلاه هو الاجراء If: ويستفاد منه في حساب متغير جديد ولكن لمجموعة المشاهدات التي ينطبق عليها الشرط. ويشمل هذا الامر الفرعي الابعازات التالية: يساوي EQ، لا يساوي NE، اقل من LT، اكثر من GT، اقل من أو يساوي LE، اكثر من أو يساوي GE. كان يكون الابعاز مثلا $X_1 = X_2$ فان $X_1 = X_2$.

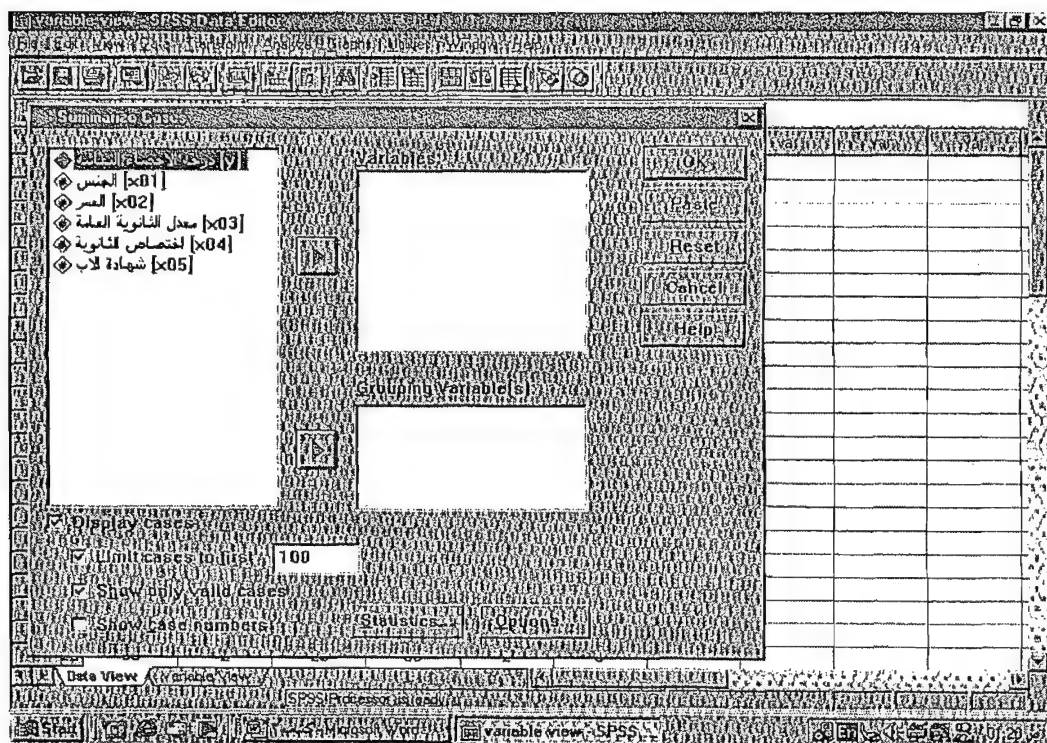
3.3. التوزيع التكراري. Frequency. باستخدام برنامج SPSS.

ان توزيع المشاهدات (التكرارات) على المتغيرات التي تم ادخالها في الفقرة "2-3" اعلاه يمكن ان يتم من خلال عدة اوامر فرعية، الا ان اكثر الطرق فعالية وتفصيلا في مجال تبويب البيانات هي الامر الفرعي Case summaries لحالة الحصول على مخرجات مجملية، في حين باستخدام الامر الفرعي Frequencies يمكن الحصول على تفاصيل اكثر واشمل. ولتشابه مسار الاجراءات المطلوبة لكلا الحالتين، سنتابع فيما يلي كيفية الحصول على مخرجات الامر الفرعي Case summaries ونعرض بعد ذلك نموذجا لمخرجات الامر الفرعي الآخر Frequency.

ويساعدنا استخدام هذه الطرق ايضا على تدقيق البيانات التي تم ادخالها، فقد يحصل وقوع اخطاء خلال عملية الادخال كأن نسجل الرقم 10 بدلا من 01 مثلا أو نعطي رمز الذكر بدلا من الانثى أو العكس، فاذا كان مثلا عدد الذكور 16 وعدد الاناث 15 وجاءت نتيجة التبويب باستخدام الامر بان عدد الذكر 17 وعدد الاناث 14 فسنكتشف ان احدى مشاهدات الاناث قد تم اعطاؤها رمز الذكور وهكذا. كما ان الامر الفرعي Statistics سيحدد العديد من مقاييس التحليل الوصفي للبيانات الخاضعة للتحليل، منها مقاييس النزعة المركزية (المتوسطات)، ومقاييس التشتت ونسب ما يشكله كل متغير، بالاضافة الى نتائج عديدة اخرى تتعلق بشكل توزيع البيانات وتباينها منها مقاييس الالتواء Skewness التي تعبر عن اتجاه ميل التوزيع التكراري ودرجته وعن تفرطحه أو درجة تدبده Kurtosis.

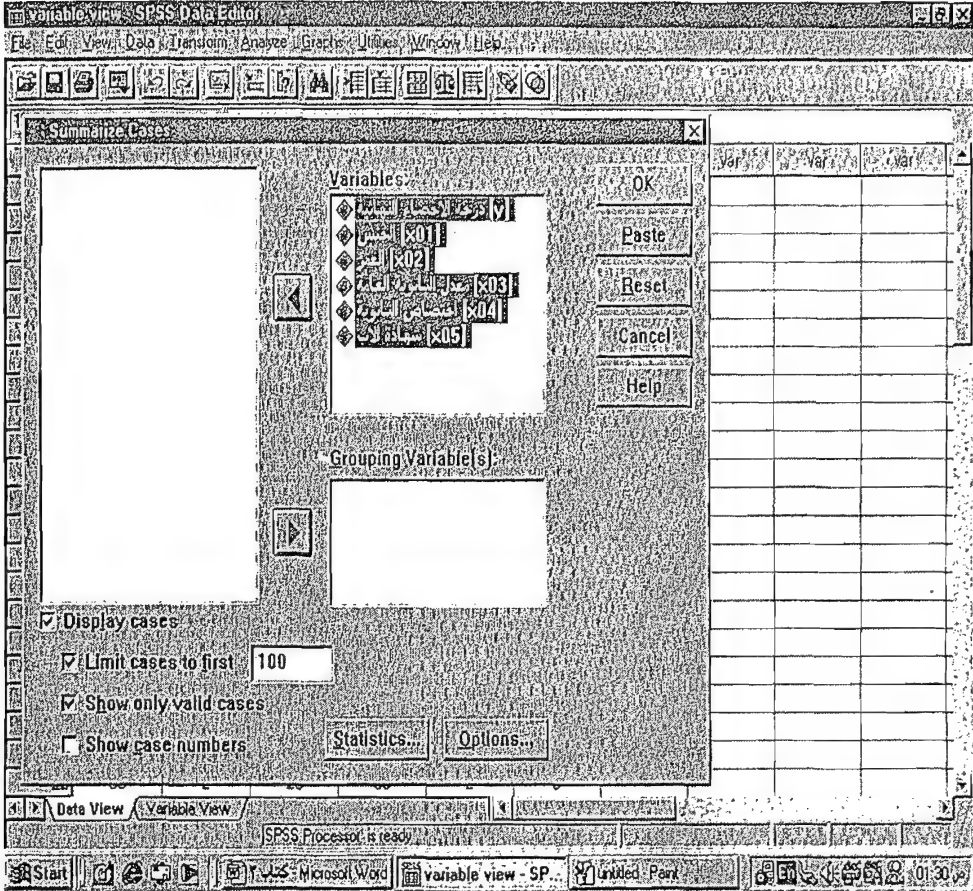
فبالنسبة للوصول الى مخرجات Case summaries يتم عبر الخطوات التالية:
 Analyze → Report → Case summaries
 في الشكل رقم (3-4)، فيتم تحديد (تضليل) المتغيرات المطلوب توزيع المشاهدات عليها، وبالكبس على السهم الموجود الى يمين المتغيرات يتم انتقال المتغيرات الى الجزء الايمن من مربع الحوار والذي يحمل عنوان متغيرات كما مبين في الشكل البياني رقم (3.5).

الشكل البياني رقم (3.4)
 يوضع مربع الحوار للاسرة الفرعي Case summaries



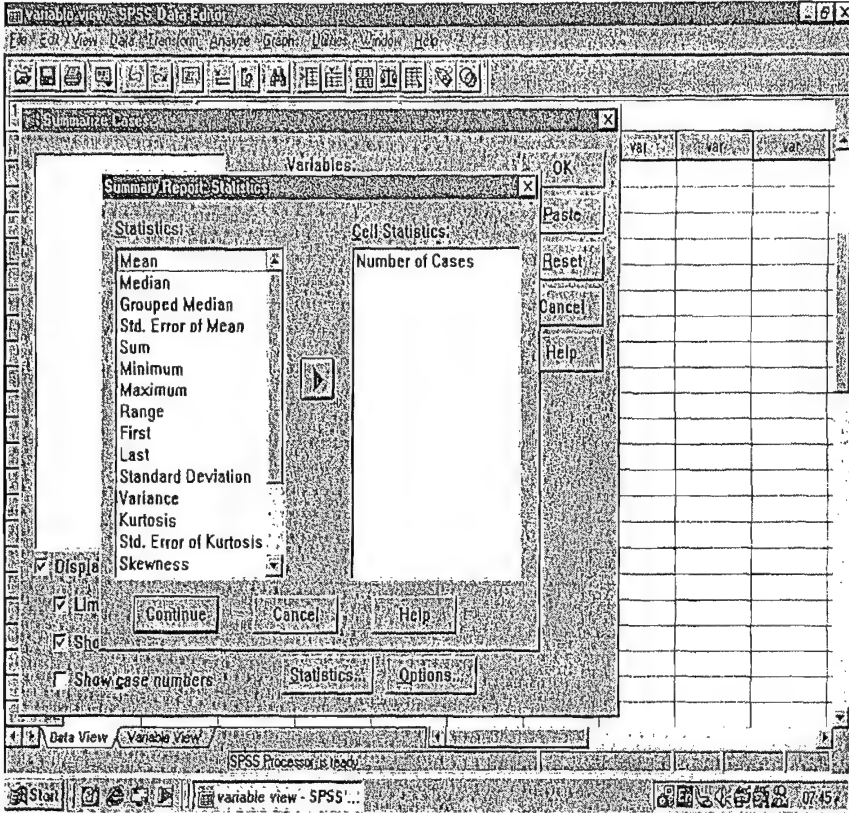
الشكل البياني رقم (5.3)

يوضع موقع التفرعات المطلوب إخضاعها للتحليل باستخدام Case summaries



ثم يتم الكبس على ايقونة Statistics لنحصل على مربع حوار آخر المبين في الشكل رقم (6.3) ليتم فيه اختيار المقاييس والمؤشرات الاحصائية المراد الحصول عليها ضمن المخرجات.

الشكل البياني رقم (6.3)
 يوضع مربع حوار المقاييس الاحصائية لفقرة Statistics



وبعد اختيار المقاييس الاحصائية المطلوبة يتم الكبس على ايقونة continue للعودة الى مربع الحوار الاول (الشكل البياني 5.3)، وبنفس الطريقة يمكن الكبس على ايقونة option لاجل تدوين العنوان المطلوب ان تحمله المخرجات ومن ثم الرجوع الى مربع الحوار الاول ايضا، والان المطلوب هو الكبس على ايقونة ok لاجراء عملية التحليل وظهور المخرجات المبينة في الجدول رقم (1.3).

المجدول رقم (1.3)

مخرجات الامر الفرعي Case summaries لتوزيع التكرارات والحصول على مقاييس
الفرعة المركزية ومقاييس التشتت ودرجة الالتواء والتفرطع.

| | Cases | | | | | |
|-----------------------|----------|---------|----------|---------|-------|---------|
| | Included | | Excluded | | Total | |
| | N | Percent | N | Percent | N | Percent |
| درجة الإحصاء النهائية | 31 | 100.0% | 0 | .0% | 31 | 100.0% |
| الجنس | 31 | 100.0% | 0 | .0% | 31 | 100.0% |
| العمر | 31 | 100.0% | 0 | .0% | 31 | 100.0% |
| معدل الثانوية العامة | 31 | 100.0% | 0 | .0% | 31 | 100.0% |
| اختصاص الثانوية | 31 | 100.0% | 0 | .0% | 31 | 100.0% |
| شهادة الأب | 31 | 100.0% | 0 | .0% | 31 | 100.0% |

Case Summaries

| | | درجة الإحصاء النهائية | الجنس | العمر | معدل الثانوية العامة | الخصائص الثانوية | شهادة الأب |
|-------|-----------------------|--------------------------|--------|--------|-------------------------|---------------------|------------|
| 1 | | 41 | 1 | 20 | 61 | 1 | 3 |
| 2 | | 40 | 2 | 22 | 70 | 2 | 4 |
| 3 | | 91 | 2 | 21 | 71 | 2 | 4 |
| 4 | | 75 | 2 | 23 | 69 | 1 | 3 |
| 5 | | 75 | 1 | 20 | 65 | 2 | 3 |
| 6 | | 64 | 1 | 22 | 59 | 1 | 3 |
| 7 | | 58 | 1 | 22 | 59 | 1 | 3 |
| 8 | | 42 | 2 | 2 | 56 | 1 | 3 |
| 9 | | 56 | 1 | 23 | 60 | 1 | 3 |
| 10 | | 52 | 2 | 24 | 65 | 2 | 3 |
| 11 | | 50 | 2 | 20 | 68 | 1 | 3 |
| 12 | | 95 | 1 | 21 | 78 | 2 | 5 |
| 13 | | 61 | 1 | 23 | 72 | 2 | 4 |
| 14 | | 68 | 2 | 33 | 65 | 1 | 4 |
| 15 | | 63 | 1 | 20 | 59 | 2 | 4 |
| 16 | | 65 | 2 | 25 | 62 | 2 | 1 |
| 17 | | 68 | 2 | 22 | 60 | 1 | 2 |
| 18 | | 70 | 1 | 23 | 72 | 1 | 5 |
| 19 | | 60 | 1 | 22 | 70 | 2 | 4 |
| 20 | | 83 | 1 | 21 | 80 | 2 | 4 |
| 21 | | 84 | 1 | 20 | 81 | 2 | 4 |
| 22 | | 88 | 2 | 20 | 83 | 2 | 3 |
| 23 | | 51 | 1 | 20 | 55 | 1 | 3 |
| 24 | | 73 | 1 | 23 | 58 | 1 | 3 |
| 25 | | 75 | 2 | 21 | 61 | 1 | 2 |
| 26 | | 79 | 2 | 23 | 67 | 1 | 3 |
| 27 | | 80 | 1 | 23 | 69 | 2 | 4 |
| 28 | | 67 | 1 | 22 | 60 | 2 | 4 |
| 29 | | 63 | 1 | 24 | 58 | 1 | 4 |
| 30 | | 66 | 2 | 21 | 62 | 1 | 3 |
| 31 | | 51 | 2 | 22 | 57 | 1 | 3 |
| Total | Minimum | 40 | 1 | 2 | 55 | 1 | 1 |
| | Maximum | 95 | 2 | 33 | 83 | 2 | 5 |
| | Range | 55 | 1 | 31 | 28 | 1 | 4 |
| | Std.Deviation | 14.46 | .51 | 4.37 | 7.68 | .51 | .84 |
| | Kurtosis | -.522 | -2.098 | 15.012 | -.225 | -2.098 | 1.145 |
| | Skewness | .019 | .204 | -2.528 | .768 | .204 | -.413 |
| | Std.Error of Kurtosis | .821 | .821 | .821 | .821 | .821 | .821 |
| | Std.Error of Skewness | .421 | .421 | .421 | .421 | .421 | .421 |
| | Geometric Mean | 64.66 | 1.37 | 20.44 | 65.13 | 1.37 | 3.23 |
| | Harmonic Mean | 63.01 | 1.29 | 16.63 | 64.73 | 1.29 | 3.05 |
| | Mean | 66.26 | 1.45 | 21.55 | 65.55 | 1.45 | 3.35 |
| | Median | 66.00 | 1.00 | 22.00 | 65.00 | 1.00 | 3.00 |
| | Grouped Median | 66.00 | 1.45 | 21.83 | 63.80 | 1.45 | 3.38 |

a. Limited to first 100 cases.

اما في حالة اختيار الطريقة Frequency من الامر الفرعي Descriptive Statistics → Analyze → Descriptive Statistics → Frequency اي: وبتوظيف ذات البيانات موضوع المثال (1.3)، فان شكل المخرجات التي سنحصل عليها سيكون كما هو مبين في الجدول رقم (2.3).

جدول رقم (2.3): مخرجات طريقة Frequency

العمر

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|---------|---------------|--------------------|
| Valid 2 | 1 | 3.2 | 3.2 | 3.2 |
| 20 | 7 | 22.6 | 22.6 | 25.8 |
| 21 | 5 | 16.1 | 16.1 | 41.9 |
| 22 | 7 | 22.6 | 22.6 | 64.5 |
| 23 | 7 | 22.6 | 22.6 | 87.1 |
| 24 | 2 | 6.5 | 6.5 | 93.5 |
| 25 | 1 | 3.2 | 3.2 | 96.8 |
| 33 | 1 | 3.2 | 3.2 | 100.0 |
| Total | 31 | 100.0 | 100.0 | |

الجنس

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|---------|---------------|--------------------|
| Valid 1 | 17 | 54.8 | 54.8 | 54.8 |
| 2 | 14 | 45.2 | 45.2 | 100.0 |
| Total | 31 | 100.0 | 100.0 | |

درجة الإحصاء النهائية

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|--------------|-----------|--------------|---------------|--------------------|
| Valid 40 | 1 | 3.2 | 3.2 | 3.2 |
| 41 | 1 | 3.2 | 3.2 | 6.5 |
| 42 | 1 | 3.2 | 3.2 | 9.7 |
| 50 | 1 | 3.2 | 3.2 | 12.9 |
| 51 | 2 | 6.5 | 6.5 | 19.4 |
| 52 | 1 | 3.2 | 3.2 | 22.6 |
| 56 | 1 | 3.2 | 3.2 | 25.8 |
| 58 | 1 | 3.2 | 3.2 | 29.0 |
| 60 | 1 | 3.2 | 3.2 | 32.3 |
| 61 | 1 | 3.2 | 3.2 | 35.5 |
| 63 | 2 | 6.5 | 6.5 | 41.9 |
| 64 | 1 | 3.2 | 3.2 | 45.2 |
| 65 | 1 | 3.2 | 3.2 | 48.4 |
| 66 | 1 | 3.2 | 3.2 | 51.6 |
| 67 | 1 | 3.2 | 3.2 | 54.8 |
| 68 | 2 | 6.5 | 6.5 | 61.3 |
| 70 | 1 | 3.2 | 3.2 | 64.5 |
| 73 | 1 | 3.2 | 3.2 | 67.7 |
| 75 | 3 | 9.7 | 9.7 | 77.4 |
| 79 | 1 | 3.2 | 3.2 | 80.6 |
| 80 | 1 | 3.2 | 3.2 | 83.9 |
| 83 | 1 | 3.2 | 3.2 | 87.1 |
| 84 | 1 | 3.2 | 3.2 | 90.3 |
| 88 | 1 | 3.2 | 3.2 | 93.5 |
| 91 | 1 | 3.2 | 3.2 | 96.8 |
| 95 | 1 | 3.2 | 3.2 | 100.0 |
| Total | 31 | 100.0 | 100.0 | |

معدل الثانوية العامة

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|----|-----------|---------|---------------|--------------------|
| Valid | 55 | 1 | 3.2 | 3.2 | 3.2 |
| | 56 | 1 | 3.2 | 3.2 | 6.5 |
| | 57 | 1 | 3.2 | 3.2 | 9.7 |
| | 58 | 2 | 6.5 | 6.5 | 16.1 |
| | 59 | 3 | 9.7 | 9.7 | 25.8 |
| | 60 | 3 | 9.7 | 9.7 | 35.5 |
| | 61 | 2 | 6.5 | 6.5 | 41.9 |
| | 62 | 2 | 6.5 | 6.5 | 48.4 |
| | 65 | 3 | 9.7 | 9.7 | 58.1 |
| | 67 | 1 | 3.2 | 3.2 | 61.3 |
| | 68 | 1 | 3.2 | 3.2 | 64.5 |
| | 69 | 2 | 6.5 | 6.5 | 71.0 |
| | 70 | 2 | 6.5 | 6.5 | 77.4 |
| | 71 | 1 | 3.2 | 3.2 | 80.6 |
| | 72 | 2 | 6.5 | 6.5 | 87.1 |
| | 78 | 1 | 3.2 | 3.2 | 90.3 |
| | 80 | 1 | 3.2 | 3.2 | 93.5 |
| | 81 | 1 | 3.2 | 3.2 | 96.8 |
| | 83 | 1 | 3.2 | 3.2 | 100.0 |
| Total | | 31 | 100.0 | 100.0 | |

اختصاص الثانوية

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|-------|---|-----------|---------|---------------|--------------------|
| Valid | 1 | 17 | 54.8 | 54.8 | 54.8 |
| | 2 | 14 | 45.2 | 45.2 | 100.0 |
| Total | | 31 | 100.0 | 100.0 | |

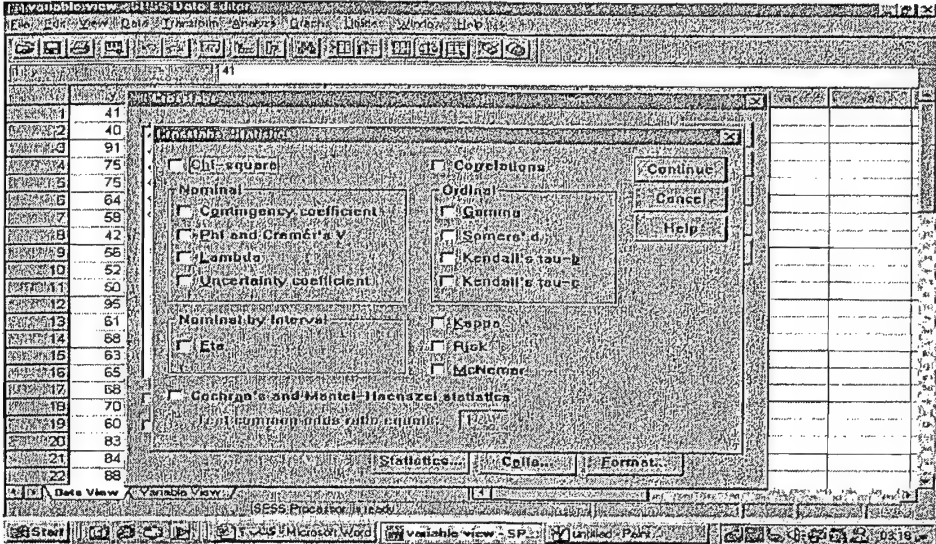
شهادة الأب

| | Frequency | Percent | Valid Percent | Cumulative Percent |
|---------|-----------|---------|---------------|--------------------|
| Valid 1 | 1 | 3.2 | 3.2 | 3.2 |
| 2 | 2 | 6.5 | 6.5 | 9.7 |
| 3 | 15 | 48.4 | 48.4 | 58.1 |
| 4 | 11 | 35.5 | 35.5 | 93.5 |
| 5 | 2 | 6.5 | 6.5 | 100.0 |
| Total | 31 | 100.0 | 100.0 | |

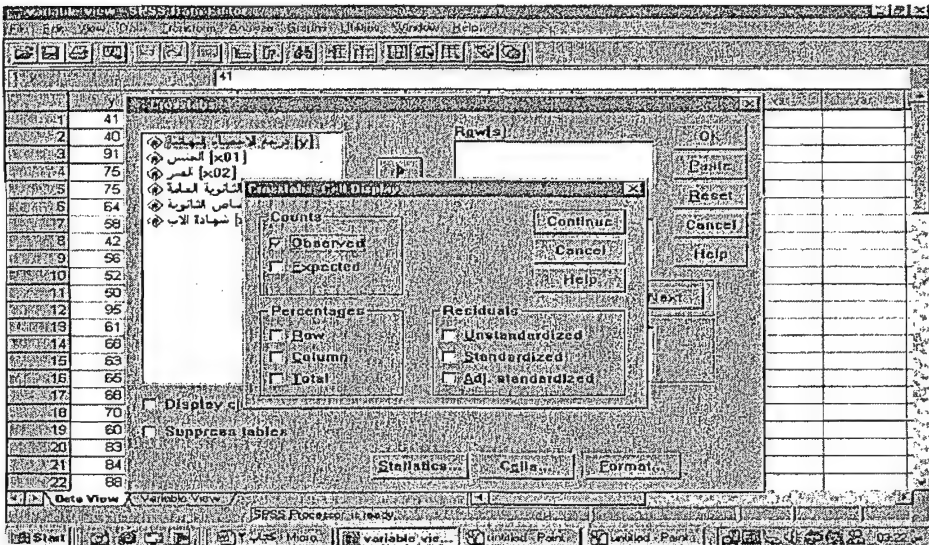
4.3 التوزيع التكراري المتعدد Cross tabs باستخدام برنامج SPSS

ويستخدم هذا النوع من التحليل لعرض تبويب متغيرين أو أكثر، مما يساعد على معرفة مدى تأثير متغير ما على متغير آخر، كمعرفة مدى علاقة مثلاً معدل الثانوية العامة على مستوى أداء الطالب في الجامعة، وذلك من خلال الحصول على نسبة معدلات الطلبة في الثانوية العامة اتجاه متغير الاداء، فإذا ما ظهر أن هناك نسبة عالية من ذوي المعدلات العالية في الثانوية العامة في خانة الاداء العالي مثلاً فإننا نستدل على أن مستوى الاداء يزداد بارتفاع معدل الثانوية العامة وهكذا. وهذا بدوره يدلنا على اتجاه العلاقة أن كانت سالبة أو موجبة. كما ويتيح لنا مربع الحوار المتعلق بـ Statistics المبين في الشكل البياني رقم (7.3) الحصول على مقاييس اختبار مربعات كاي Chi Square والمعامل التوافقي Contingency Coefficient و Lambda ومعامل الارتباط Correlations وغيرها. كما ويتيح المربع الآخر في الامر والمتعلق بـ Cell Display المبين في الشكل البياني رقم (8.3) الحصول على النسب والقيم المعيارية Standardization.

الشكل البياني رقم (7.3):
 بوضع سريع حوار Statistics للامر Crosstabs



الشكل البياني رقم (8.3):
 بوضع مربع الحوار Cell Display للامر Crosstabs



والوصول الى استخدام هذه الطريقة يأخذ المسار التالي :

Analyze → Descriptive Statistics → Crosstabs

و باخضاع البيانات موضوع المثال (1.3) للامر Cross tabs بعد المرور بمربعات الحوار نحصل على جداول المخرجات وعددها 16 جدولاً تعود لخمس متغيرات موزعة على المتغير التابع Dependent Variable وهو متغير علامات الاحصاء النهائية، ونختار من بينها للعرض والتحليل مخرجات متغير واحد كنموذج وهو متغير الجنس (ذكور، اناث) باعتباره المتغير الاول في القائمة من حيث التسلسل والمبينة نتائج في جدول رقم (3.3) التالي:

جدول رقم (3.3)

مخرجات الامر Crosstabs لتحليل بيانات المثال (1.3)

Case Processing Summary

| | Cases | | | | | |
|--|-------|---------|---------|---------|-------|---------|
| | Valid | | Missing | | Total | |
| | N | Percent | N | Percent | N | Percent |
| الجنس * درجة الإحصاء النهائية | 31 | 100.0 | 0 | .0% | 31 | 100.0% |
| العمر * درجة الإحصاء النهائية | 31 | 100.0 | 0 | .0% | 31 | 100.0% |
| معدل الثانوية العامة * درجة الإحصاء النهائية | 31 | 100.0 | 0 | .0% | 31 | 100.0% |
| اختصاص الثانوية * درجة الإحصاء النهائية | 31 | 100.0 | 0 | .0% | 31 | 100.0% |
| شهادة الأب * درجة الإحصاء النهائية | 31 | | 0 | .0% | 31 | 100.0% |

الجنس * درجة الإحصاء النهائية

Crosstab

| | | درجة الإحصاء النهائية | | | | | | | | | |
|---------|-------------------------------|-----------------------|--------|--------|--------|--------|--------|--------|--------|--------|-----|
| | | 40 | 41 | 42 | 50 | 51 | 52 | 56 | 58 | 60 | ▶ |
| الجنس 1 | Count | | 1 | | | 1 | | 1 | 1 | 1 | ... |
| | %Within الجنس | | 5.9% | | | 5.9% | | 5.9% | 5.9% | 5.9% | ... |
| | درجة الإحصاء النهائية %Within | | 100.0% | | | 50.0% | | 100.0% | 100.0% | 100.0% | ... |
| 2 | Count | 1 | | | | | | | | | ... |
| | %Within الجنس | 7.1% | | 7.1% | 7.1% | 7.1% | 7.1% | | | | ... |
| | درجة الإحصاء النهائية %Within | 100.0% | | 100.0% | 100.0% | 50.0% | 100.0% | | | | ... |
| | Count | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | ... |
| | %Within الجنس | 3.2% | 3.2% | 3.2% | 3.2% | 6.5% | 3.2% | 3.2% | 3.2% | 3.2% | ... |
| | درجة الإحصاء النهائية %Within | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | 100.0% | ... |
| Total | | | | | | | | | | | |

Chi-Square Tests

| | Value | Df | Asymp.Sig. (2-sided) |
|-------------------------------|-----------------------|----|-------------------------|
| Pearson Chi-Square | 26.289 ^(a) | 25 | .392 |
| Likelihood Ratio | 36.093 | 25 | .070 |
| Linear-by- Linear Association | .193 | 1 | .660 |
| N of Vaild Cases | 31 | | |

Symmetric Measures

| | Value | Asymp.Std. Error ^(b) | Approx. T ^(c) | Approx. Sig. ^(d) |
|--|-------|------------------------------------|-----------------------------|-----------------------------------|
| Interval by interv Pearson's R | -.080 | .180 | -.434 | .668 ^(d) |
| Oridunal Ordin Spearman Correlation | -.040 | .186 | -.215 | .831 ^(d) |
| N of Vaild Cases | 31 | | | |

^(a) 52 Cells (100.0%) have expected count less than 5. The minimum expected count is .45.

^(b) Not assuming the null hypothesis.

^(c) Using the asymptotic standard error assuming the null hypothesis.

^(d) Based on normal approximation.

5.3 تفسير مخرجات Cross tabs

ومن مخرجات متغير الجنس موزعا على المتغير التابع (علامات الاحصاء النهائية) الواردة في الجداول اعلاه نستدل على مايلي :

الجدول الاول: الاستدلال على اكتمال كافة المشاهدات لكافة المتغيرات وكما تشير لذلك النسب 100%، وبالتالي فان نسبة القيم المفقودة هي 0%.

الجدول الثاني: ان 66.7% من عدد الطالبات الاناث هم ضمن علامة 75% . في حين المتوقع وفقا للتوزيع النظري Expected ان تكون النسبة بحدود 10%.

الجدول الثالث: لم تدل نتائج اختبار Chi Square على تجانس معنوي في توزيع الاناث وفقا لفئات العلامات حيث في الغالب كانت علاماتهم متركزة في الفئات بعد المتوسط، وكما يتبين ذلك من مستوى الدلالة (المعنوية) في جدول المخرجات.

الجدول الرابع: ان درجة التماثل في شكل التوزيع الطبيعي Normal بموجب معيار Lambda هو عالي المعنوي (عند $\alpha = 0.000$) حيث بلغت قيمة Lambda ما مقداره 0.857 (من مجموع 1)، اما المعايير المتعلقة بدرجة الارتباط فقد جاءت ضعيفة نسبيا سواء بموجب معيار Pearson R أو Spearman correlation أو كما تم الاستدلال مسبقا من Chi Square بالنسبة لمعيار معامل الارتباط التوافقي Contingency Coefficient .

5.3 توزيع التكرارات على فئات باستخدام برنامج EXCEL

ونتابع اجراء العمليات التحليلية المتعلقة بتبويب البيانات في فئات تكرارية وكذلك عرض البيانات باستخدام الحاسوب ولكن هذه المرة باستخدام برنامج EXCEL وستشمل الفقرات التالية كلاً من التوزيع التكراري على شكل فئات والعرض البياني، حيث يمكن اعتبار هذا البرنامج في هذه الحالة هو الاكثر سهولة والاقل حاجة للوقت كما اسلفنا في اعلاه. اما في حالة الرغبة في استخدام برنامج SPSS لاغراض العرض البياني فالاجراءات المطلوبة هي اختيار الامر الرئيسي Graphics ومن ثم

تحديد نوع الرسمة المطلوبة، من خلال تحديد المتغيرات المطلوب رسمها، وعلى افتراض توفر ملف بهذه المتغيرات.

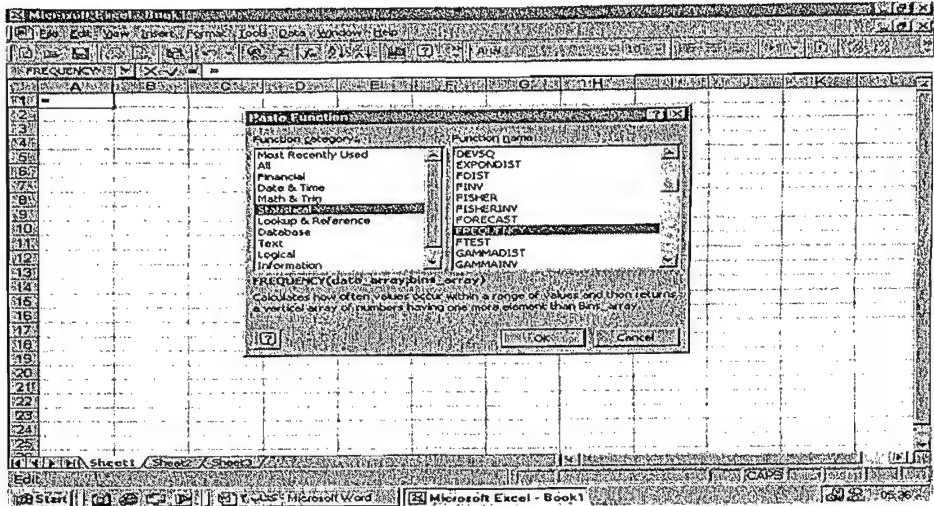
والمقصود بالفئات هنا هو تقسيم البيانات الى مجموعات تدعى "فئات Intervals"، فلو افترضنا ان المطلوب هو توزيع الطلبة على فئات درجات مادة الاحصاء، وان عدد الطلبة البالغ 31 طالبا والفئات التي يوزع عليها الطلبة هي كما مبين في ادناه:

درجات مادة احصاء: 41، 43، 91، 75، 76، 64، 60، 40، 56، 52، 50، 95، 61، 68، 63، 65، 68، 70، 60، 83، 84، 88، 51، 73، 75، 79، 80، 67، 63، 66، 51.
الفئات: 40-49، 50-59، 60-69، 70-79، 80-89، 90-99.

مع الاشارة الى انه في حالة لدينا احد حدي الفئة هو اقل أو اكثر فيتم عادة اعتماد طول الفئة السابقة لها أو الفئة اللاحقة لتعيين الحد الاعلى أو الادنى غير الموجود، وللزيادة في التفصيل النظري وفي كيفية تحديد عدد الفئات يمكن الرجوع الى فقرة " الطريقة اليدوية " التي سيتم التطرق اليها لاحقا. ان الاجراءات المطلوبة لانجاز عملية التوزيع التكراري الى فئات باستخدام برنامج EXCEL هي كما يلي:

- الدخول على البرنامج من خلال: Excel → programs → start .
- بعد ظهور صفحة البرنامج نكبس (click) على معالج الدوال fx الموجود في اعلى الصفحة أو الحصول عليه من الامر الرئيسي "دراج Insert" فنحصل على مربع الحوار المبين في الشكل (9.3) في ادناه:

شكل بياني رقم (9.3):
موضوع قائمة معالج الدوال لبرنامج Excel

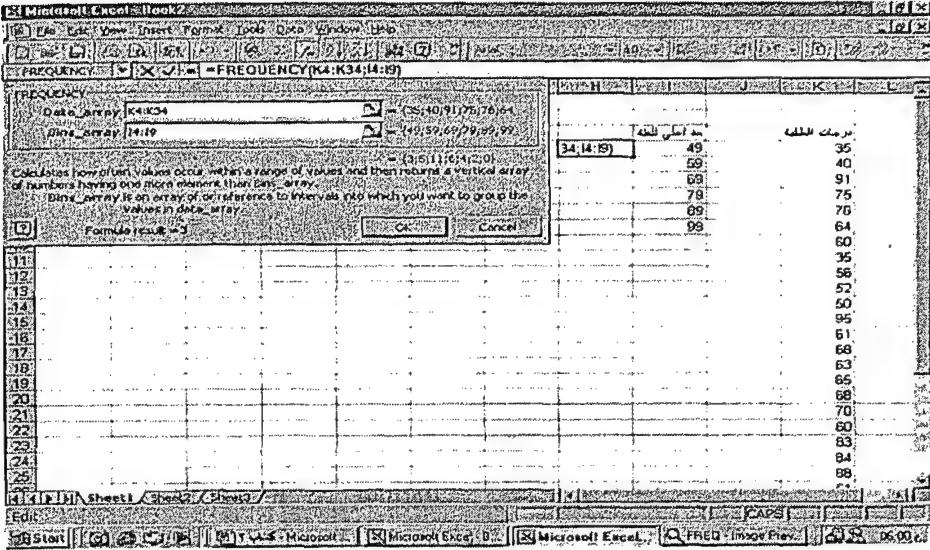


- نقوم بالتأشير على فئة statistical وكذلك على الدالة أو الوظيفة المطلوبة وهي Frequency الواقعة في الجانب الايمن من مربع الحوار كما هي مبينة في الشكل اعلاه، ثم نكبس (click) على موافق (ok) الموجودة في اسفل المربع، فتظهر اشروطة الدالة ليتم فيها تعبئة البيانات المطلوبة والمتمثلة بدرجات الطلبة في مادة الاحصاء، وفي الشريط الآخر الحدود العليا للفئات فنحصل على النتيجة في الجدول ادناه رقم (4.3) والمبينة في الشكل البياني رقم (10.3).

جدول رقم (4.3): جدول التوزيع التكراري لفئات درجات الطلبة

| الفئات | التكرار |
|---------|-----------------|
| 40 - 49 | 3 |
| 50 - 59 | 5 |
| 60 - 69 | 11 |
| 70 - 79 | 6 |
| 80 - 89 | 4 |
| 90 - 99 | 2 |
| المجموع | $\sum f_i = 31$ |

الشكل رقم (10.3): شريط مدخلات الدالة Frequency



7.3. الرسوم والاشكال البيانية باستخدام برنامجي

EXCEL و SPSS

وهي إحدى طرق عرض البيانات التي تساعد على توضيح المعلومات الرقمية، وتعتبر أكثر فعالية في وصول مضمونها إلى القارئ بسهولة فهمها. ومنها ما تخص البيانات المبوبة المنظمة في جداول وفقاً لفئات تكرارية، وقد تكون على شكل مدرج تكراري (اعمدة متلاصقة) أو منحنيات أو مضلعات بيانية. وأخرى تتعلق بجداول بيانات غير مبوبة مصنفة حسب صفات نوعية تبعا لطبيعة الظاهرة تحت الدراسة التي قد تكون زمنية أو جغرافية، والنوع الثاني قد يكون على شكل مستطيلات منفردة أو متعددة أو مركبة، أو على شكل دائرة بيانية تعود لقطاعات أو اجزاء مختلفة لظاهرة معينة، بحيث يمثل كل قطاع بجزء من مساحة الدائرة لتسهيل عملية المقارنة بين الاجزاء، بالإضافة إلى الرسوم والصور.

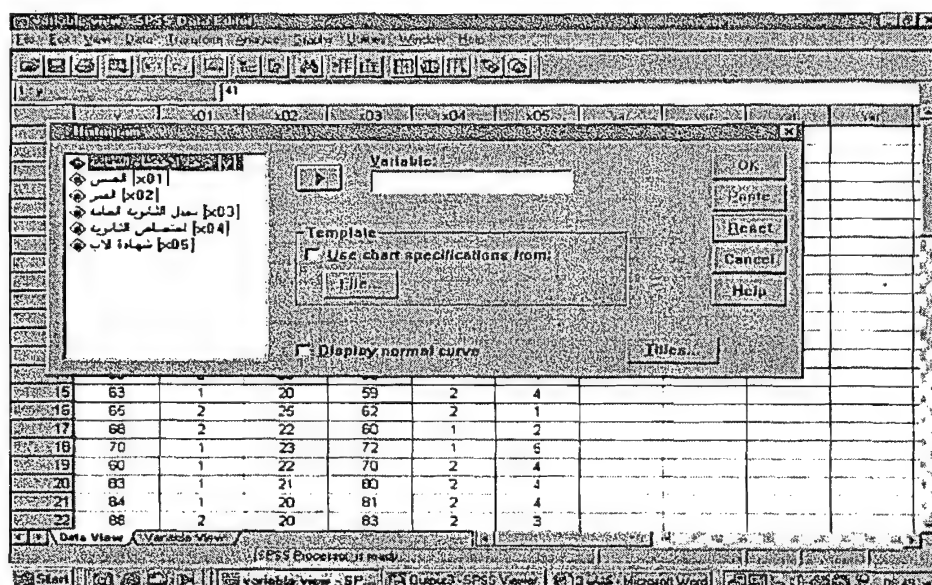
اولاً: في حالة استخدام برنامج SPSS

ان اجراءات استخدام برنامج SPSS تتلخص بالدخول الى البرنامج واختيار الامر الرئيسي Graphs ومن ثم تعيين نوع الرسم البياني المطلوب والكبس عليه (Click) للحصول على مربع الحوار ومتابعة انجاز الرسم، بافتراض أن لدينا ملفاً بالبيانات التي سنختار منها المتغير (او المتغيرات) المطلوب عرضها بيانياً، أو بناء ملف جديد والعمل عليه. مثال ذلك لو كنا بصدد عرض متغير شهادة الاب من ملف الطلبة موضوع المثال (1-3) اعلاه، واخترنا نوع الرسم المطلوب هو المدرج التكراري Histogram فستكون لدينا الخطوات التالية:

- الدخول على ملف الطلبة والتأشير على الامر الرئيسي Graphs وعند الكبس على Histogram سيظهر لنا مربع الحوار الرئيسي المبين في الشكل البياني رقم (11-3) التالي :

شكل بياني رقم (11-3)

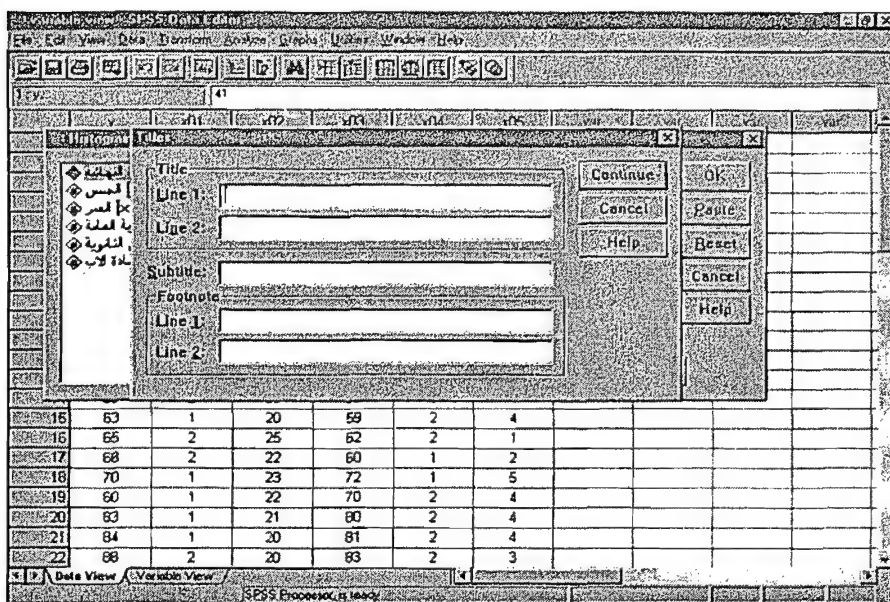
مربع الحوار الرئيسي لرسم Histogram باستخدام برنامج SPSS



- نقوم بتحويل المتغير المطلوب وهو شهادة الـ AB الى المستطيل الواقع الى اليمين بعد التأشير على المتغير واستخدام السهم، وإذا رغبتنا بظهور المنحنى الطبيعي مع المدرج نقوم بالتأشير على حقل Display Normal Curve الموجود عند اسفل مربع الحوار الرئيسي المبين في الشكل اعلاه.
- الكبس على Titles فيظهر لنا مربع الحوار الملحق المبين في الشكل البياني رقم (12-3) الذي يتم فيه ادراج عنوان الرسم.

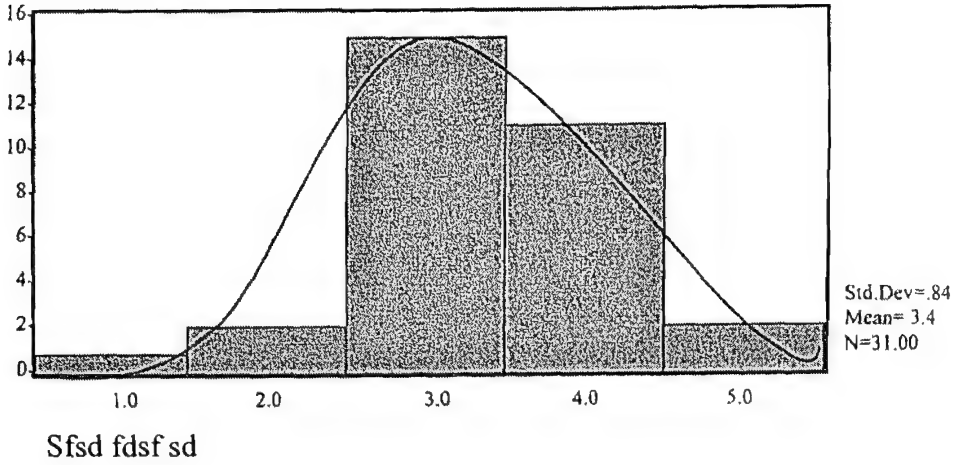
شكل بياني رقم (12-3) :

مربع الحوار الملحق بمتعلق بعنوان الرسم باستخدام برنامج SPSS



- بعد ادراج العناوين المطلوبة في مربع الحوار الملحق، نعود الى مربع الحوار الرئيسي من خلال الكبس على ايقونة Continue. وفي مربع الحوار الرئيسي يتم الكبس على ايقونة Ok فنحصل على المدرج التكراري مع المنحنى الطبيعي كما هو مبين في الشكل البياني رقم (13-3) التالي :

شكل بياني رقم (13-3)

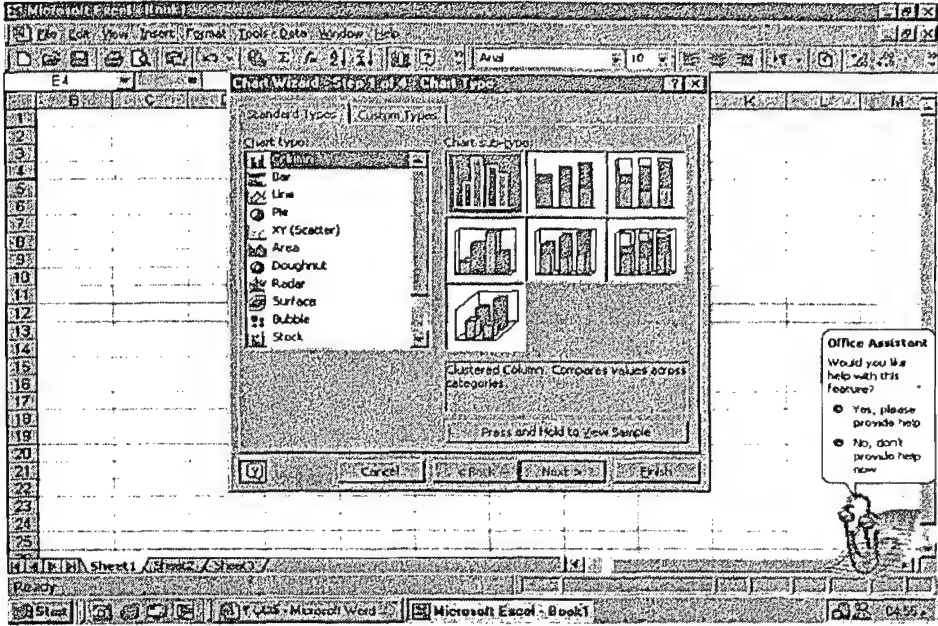


ثانياً: حالة استخدام برنامج EXCEL

اما الاجراءات المطلوبة لاستخدام برنامج EXCEL لانجاز الرسوم والاشكال البيانية فتتلخص بالخطوات التالية:

- الدخول الى البرنامج من خلال Start → Programs → Excel
- يتم اعداد جدول بالبيانات المطلوب عرضها ومن ثم تظليل الجدول أو الاجزاء المطلوب عرضها.
- يتم الكبس (click) على معالج الرسوم البيانية المتوفر على شريط الصيغ أو الحصول عليه من الامر ادراج (Insert) فتظهر لنا صفحة انواع الاشكال البيانية المبينة في الشكل البياني رقم (14-3) ادناه لنختار نوع الشكل البياني المطلوب.

الشكل البياني رقم (143)



- معاينة الشكل الذي سيبدو عليه الرسم البياني بالضغط المستمر على ايقونة to view sample المبين موقعها في الاسفل من الشكل اعلاه.
 - اختيار فئة الشكل المطلوب بالكبس على "انواع مخصصة custom types"، بعدها يتم متابعة الخيارات المتوفرة بعد الانتهاء من العمل مع كل خيار بالكبس على Next والتي تشمل :
 - تسمية سلسلة (مفاتيح) بيانات الشكل البياني بعد مرحلة اختيار فئة الشكل وظهوره.
 - ضبط الخيارات المطلوبة للشكل البياني والتي تشمل: العناوين title ووسيلة الايضاح legend والتحكم في اظهار القيم وجدول البيانات وغيرها.
 - تحديد الورقة التي نرغب بادراج الشكل البياني عليها ان كانت مع جدول البيانات أو منفصلة، وبالكبس على ايقونة Finish يتم اغلاق المعالج ويظهر الشكل البياني.
- وفي الآتي نتابع كيفية الحصول على الرسوم والأشكال البيانية كلاً على حده:

(1) المنحنيات و الخطوط البيانية التكرارية والمتجمعة

Frequency and cumulative Curve

وهي عبارة عن منحنيات متصلة تمر بجميع النقاط المحددة وفي حالة التوصيل بين هذه النقاط بخطوط مستقيمة عندها تدعى بالمضلعات التكرارية. وان خطوات تهيئة البيانات لاعداد منحنى أو مضلع تكراري تتطلب ايجاد مراكز الفئات والتي هي عبارة عن حاصل جمع حدي الفئة وقسمته على 2. اما بالنسبة للمنحنيات أو المضلعات المتجمعة الصاعدة Ascending Ogive Curve أو النازلة Descending Ogive Curve فتتطلب ايجاد قيم المتجمعات وفقا لما هو وارد في الطريقة اليدوية. ولنفترض أن المطلوب ايجاد المنحنيات التكرارية والمتجمعة لبيانات الجدول (4.3) موضوع مثالنا اعلاه فيكون لدينا الجدول رقم (5.3) التالي :

جدول رقم (5.3)

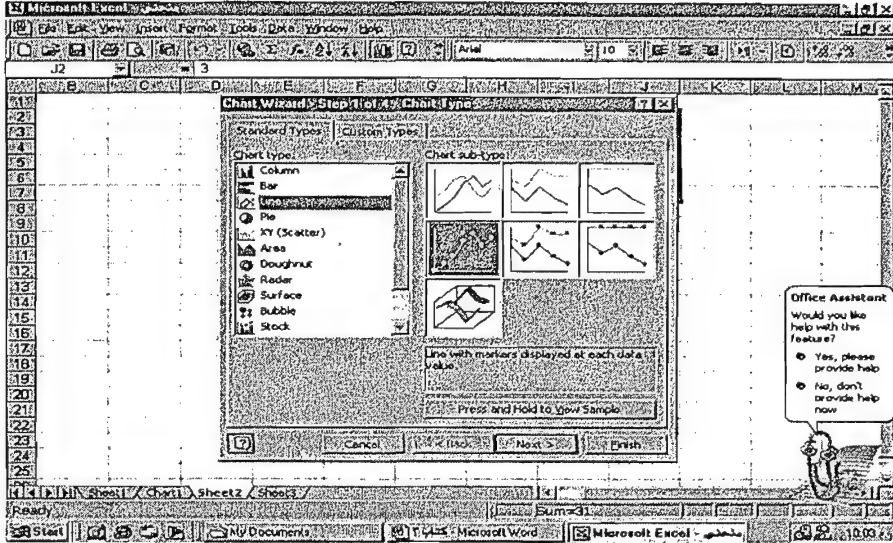
بضم الفئات التكرارية والتكرارات و مراكز الفئات والمتجمعين الصاعد والنازل

| الفئات | التكرار | المتجمع الصاعد | المتجمع النازل |
|-----------|---------|----------------|----------------|
| أقل من 50 | 3 | 3 | 31 |
| 50 – 59 | 5 | 8 | 28 |
| 60 – 69 | 11 | 19 | 23 |
| 70 – 79 | 6 | 25 | 12 |
| 80 – 89 | 4 | 29 | 6 |
| 90 فأكثر | 2 | 31 | 2 |

أ. المضلع والمنحني التكراري

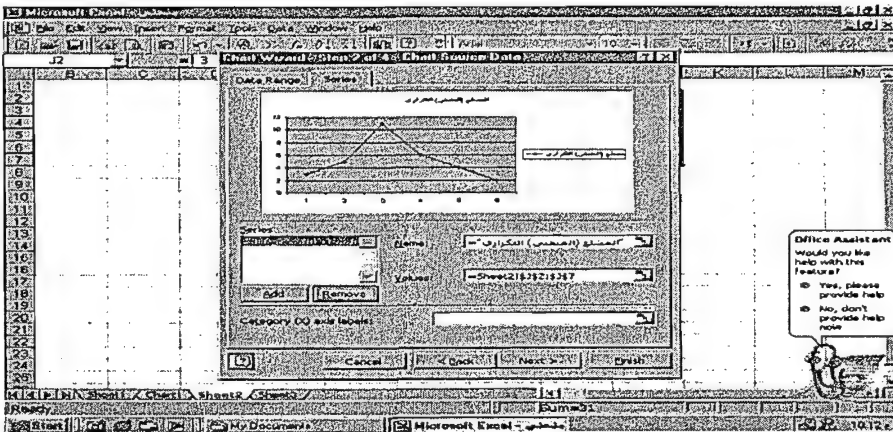
- عقب الدخول الى برنامج Excel وفق الاجراءات التي تم شرحها في اعلاه، يتم تظليل (تحديد) البيانات المطلوب عرضها كمضلع تكراري، ومن ثم يتم الكبس على معالج الرسوم البياني فنحصل على الشكل البياني رقم (14.3) ومنه نؤشر على فئة Line المبين في الشكل رقم (15.3) في ادناه :

الشكل البياني رقم (15.3) مرحلة اختيار نوع الرسم البياني



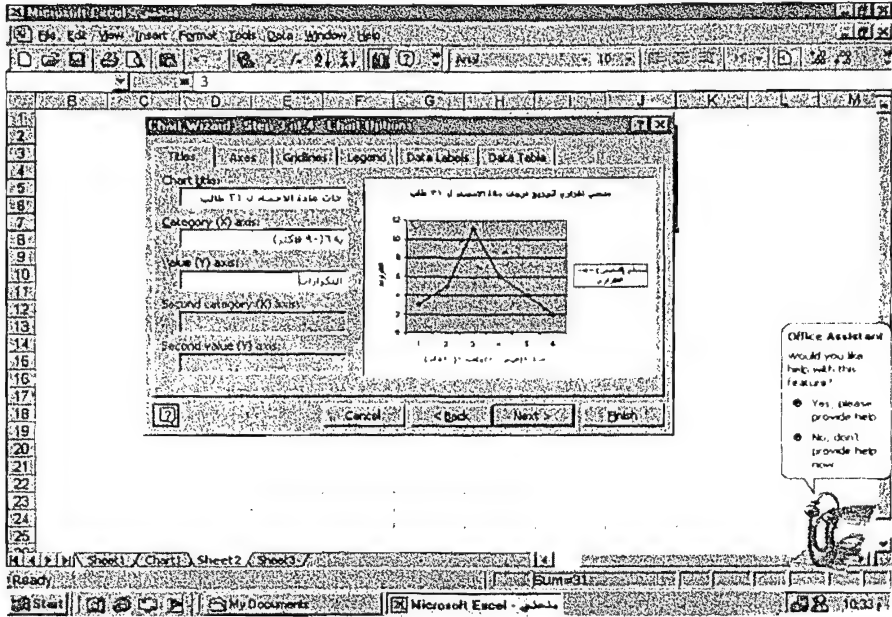
- يلي ذلك الكبس على NEXT المبينة في الشكل اعلاه، فنحصل على الشكل رقم (16.3).

الشكل رقم (16.3) المرحلة الثانية من اجراءات رسم المصطلح التكراري



- ونتابع اجراءات اكمال المنحني البياني بالكبس مرة اخرى على NEXT لنكمل تدوين عنوان الشكل البياني واسماء المحاور الأفقي (الذي عادة ما يحتوي على الفئات او مراكزها أو احد حدودها) والعمودي (الذي يتعلق بالتكرارات)، بالإضافة الى أوامر اخرى تتعلق بمظهر الشكل البياني، وكما مبين في الشكل التالي (17.3):

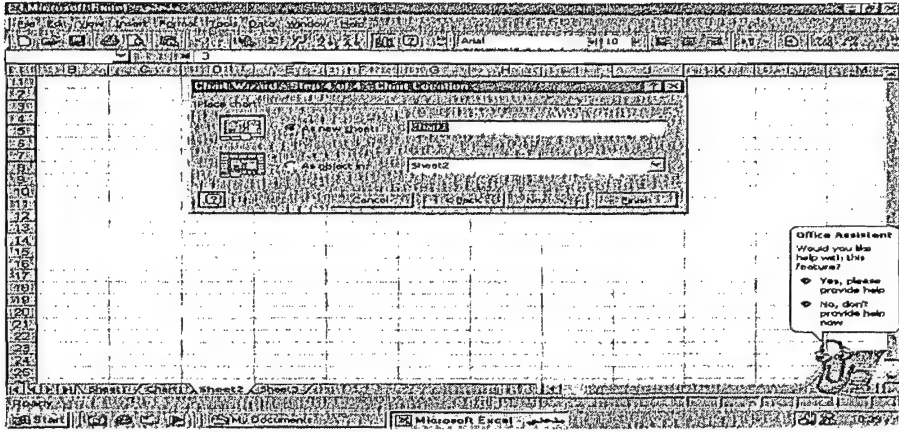
الشكل البياني (17.3) مرحلة تدوين عناوين الرسم البياني



- والكبس الاحق على NEXT يقودنا الى السؤال ان كنا نرغب باظهار الرسم على ورقة مستقلة أو بمعية جدول البيانات وكما مبين في الشكل (18.3) في ادناه:

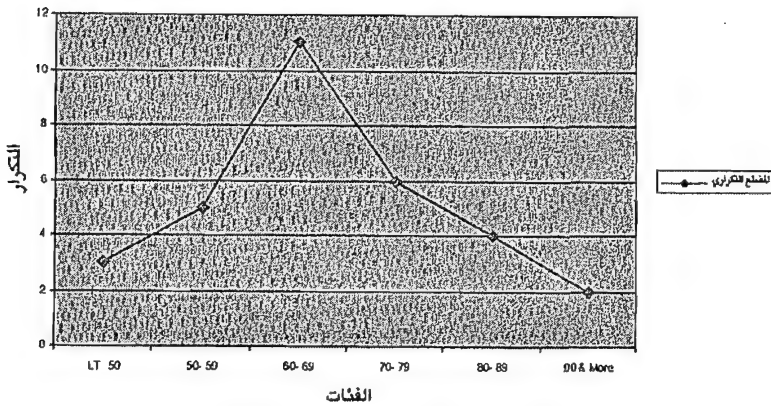
الشكل البياني (18.3)

مرحلة تحديد الورقة التي يظهر عليها الرسم البياني



- وبالكبس على ايقونة FINISH المبين موقعها على الشكل (18.3) اعلاه نحصل
رسم المنحني التكراري في الشكل (19.3) ادناه :

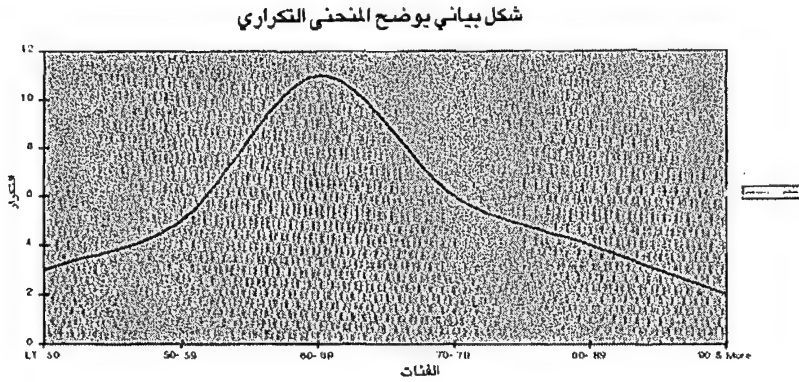
الشكل البياني رقم (19.3)



أما المنحنى التكراري فهو عبارة عن تمهيد (smoothing) لنقاط التقاء المضلعات (المستقيمات) والحصول عليه يتم باختيار منحنى (smoothing lines) بعد الكبس على ايقونة Custom Type (تخصيص) المبينة على الشكل (3-15) وبمتابعة نفس الخطوات التي تم اتباعها مع المضلع التكراري نحصل على المنحنى المبين في الشكل رقم (3-20) ادناه:

شكل بياني رقم (3-20)

شكل بياني يوضح المنحنى التكراري



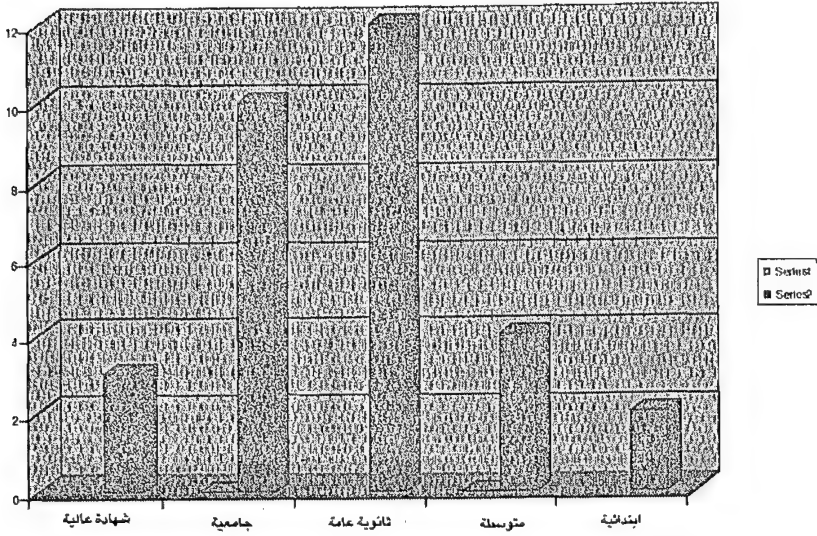
ب. المنحنى المتجمع (الصاعد والنازل)

وبتظليل البيانات المتعلقة بالمتجمع التكراري (التي يتم الحصول عليها بموجب الخطوات الموضحة في فقرة الطريقة اليدوية التي سيتم التطرق إليها لاحقاً) ومتابعة نفس الخطوات التي تم العمل بها في حالة المضلع التكراري اعلاه نحصل على رسم المنحنى المتجمع الصاعد والنازل المبين في الشكل البياني رقم (3-21) التالي:

وبالعمل بموجب الخطوات التي تطرقنا إليها في اعلاه باستثناء التأشير على اعمدة نحصل على الشكل البياني (22.3) التالي:

شكل بياني رقم (22.3)

أعمدة بيانية منفردة تعرض توزيع آباء الطلبة للشال (13) حسب الشهادة



ب. الاعمدة البيانية المتعددة والمركبة

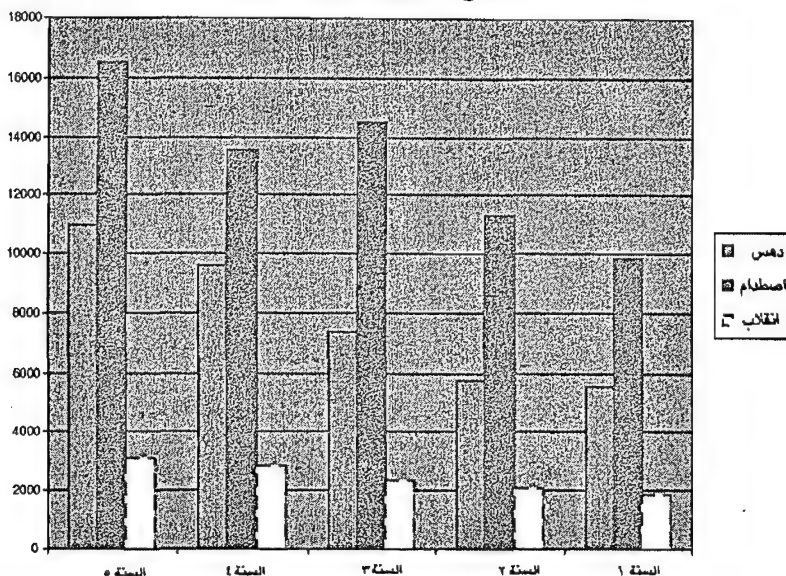
وهو الشكل البياني الذي يمكن استخدامه لعرض عدة ظواهر أو عدة مستويات للظاهرة الواحدة في عدة اعمدة، ويدعى "الاعمدة المتعددة" اما اذا تم عرض هذه الظواهر أو المستويات بذات العمود فيطلق عليه "الاعمدة المركبة" كما مبين في الاشكال البيانية (23-3) و(24-3) على التوالي والتي تعرض بيانات الجدول في ادناه.

جدول يبين عدد حوادث الطرق مصنفة حسب نوع الحادث
للسنوات الخمس الأخيرة للأحدي الدول

| السنين | السنة 1 | السنة 2 | السنة 3 | السنة 4 | السنة 5 |
|---------|---------|---------|---------|---------|---------|
| دهس | 5575 | 5764 | 7338 | 9600 | 11011 |
| اصطدام | 9865 | 11345 | 14535 | 13543 | 16554 |
| انقلاب | 1848 | 2065 | 2345 | 2865 | 3143 |
| المجموع | 17288 | 19174 | 24218 | 26008 | 30708 |

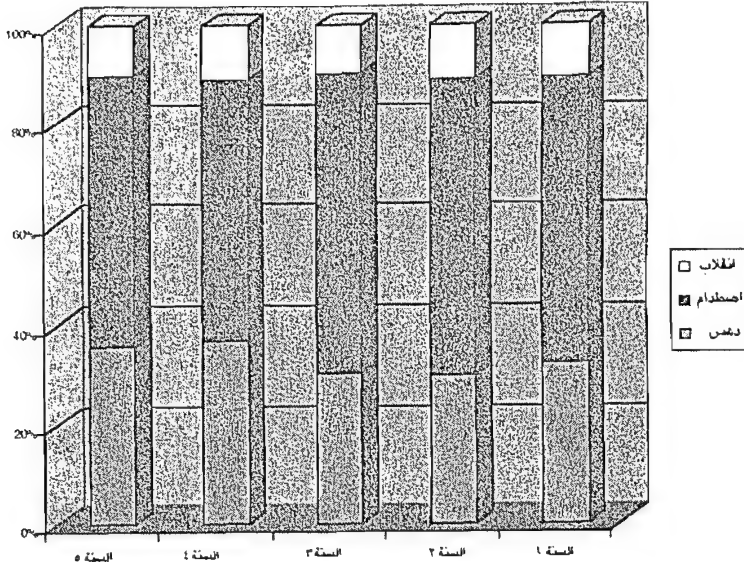
شكل بياني رقم (233)

يوضح استخدام الاعمدة المتعددة لحوادث المرور مصنفة حسب نوعها خلال السنوات
الخمس الأخيرة للأحدي الدول



شكل بياني رقم (243)

يوضح استخدام الأعمدة البيانية المركبة لعدد الحوادث مصنفة حسب نوعها خلال
الخمس سنوات الأخرى للأحدى الدول

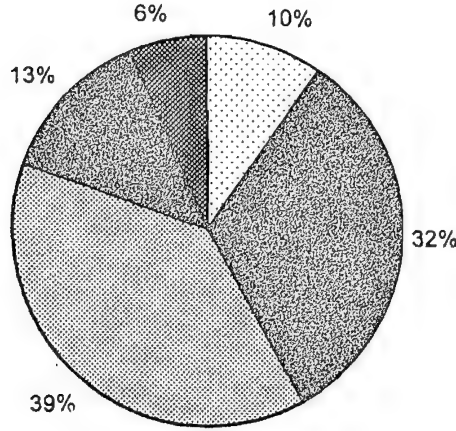


(3) الدائرة البيانية

وبتظليل (تحديد) البيانات المطلوب رسمها ولنفترض الواردة في الجدول (6.3)
واختيار النوع "الدائرة Pie" نحصل على الرسم المبين في الشكل البياني رقم (25.3)
التالي:

شكل بياني رقم (25.3)

دائرة بيانية توضح شهادة آباء عينة الطلبة: (عالية 10٪) (بكالوريوس 32٪) (ثانوية 39٪) (متوسطة 13٪) (ابتدائية 6٪)



(4) الرسوم والصور البيانية

ويمكن الاستعانة في الحصول على الرسوم والصور بالرجوع الى الامر الرئيسي ادراج Insert ومن ثم استخدام الامر الفرعي صور Picture ومنه الاجراء Clip Art، كما ويمكن الاستفادة ايضا من الامر الفرعي رموز Symbol، والاشكال رقم (3-26) نماذج لهذا النوع من العرض البياني.

اشكال بيانية رقم (263) تمثل نماذج للرسم والصور

- التعبير عن عدد الزوارق الرياضية في سواحل احدى المدن والبالغ عددها 800 زورق

$$100 \text{ زورق} = \text{صورة زورق}$$

- للتعبير عن عدد سكان احدى الدول البالغ 9 ملايين نسمة :

$$1000000 \text{ نسمة} = \text{صورة شخص}$$

- للتعبير عن عدد السيارات في احدى الدول البالغ عددها 445000 سيارة

$$100000 \text{ سيارة} = \text{صورة سيارة}$$

$$4500 \text{ سيارة} = \text{صورة سيارة}$$



8.3 الطريقة اليدوية في تبويب وعرض البيانات

1. التوزيع التكراري البسيط Simple Frequency Distribution

ان اجراءات تبويب البيانات على فئات والتي تدعى "بالفئات التكرارية"، والملاحظات التي يتم توزيعها على هذه الفئات تسمى "بالتكرارات" يتم انجازها من خلال الخطوات التالية :

الخطوة الاولى: تحديد عدد الفئات

وفيها يتم مراعاة بعض المحددات ومن اهمها طبيعة البيانات وحجمها ومقدار الاختلاف بينها، فتلك التي عددها محدود من المناسب ان يكون عدد فئاتها قليلاً، وسنحتاج الى عدد اكبر من الفئات في حالة كون عدد البيانات اكبر، ولكن بصورة عامة يجب ان لا تكون قليلة جداً بحيث نفقد عدداً من البيانات ولا نختار عدداً كبيراً من الفئات بحيث تزداد التفاصيل التي قد تكون غير مستهدفة وتؤدي الى خلق فئات خالية من التكرارات. ويمكن الاستدلال بالصيغة المقترحة من قبل (Struges, 1926) لتحديد عدد الفئات والتي صيغتها هي :

$$K=1+3.322 (\text{Log } n) \dots\dots\dots (3.1)$$

حيث إن: $k =$ عدد الفئات

$n =$ عدد القيم (البيانات)

الخطوة الثانية: إيجاد طول (او مدى) الفئة H

وهو عبارة عن الفرق بين اكبر واصغر قيمة بين البيانات وقسمته على عدد الفئات التي يتم تحديدها في الخطوة الاولى، مع محاولة تقريب النتيجة الى عدد صحيح في حالة الكسر، اي:

$$H = \frac{\text{اكبر قيمة} - \text{أصغر قيمة}}{\text{عدد الفئات}}$$

الخطوة الثالثة: تحديد حدود الفئات

فالحد الأدنى للفئة الاولى هو عبارة عن اصغر قيمة بين البيانات، والحد الاعلى سيكون عبارة عن اضافة طول الفئة الى قيمة الحد الأدنى مطروحا منه 1، والحد الأدنى الفئة الثانية هو عبارة عن القيمة اللاحقة للحد الاعلى للفئة السابقة، ثم يضاف اليه طول الفئة مطروحا منه 1 ليصبح الحد الاعلى للفئة الثانية وهكذا .

الخطوة الرابعة: توزيع التكرارات على الفئات

وفيها يتم توزيع البيانات على الفئات، وذلك بوضع اشارة امام الفئة المناسبة، ولغاية اتمام كافة البيانات ليتم بعد ذلك حساب هذه الاشارات ولكل فئة لتدوين التكرار المقابل للفئة المعنية، مع مراعاة مطابقة مجموع التكرارات لمجموع عدد البيانات.

مثال (2.3): لدينا في ادناه علامات مادة الاحصاء لعينة الطلبة البالغ عددها 31 طالباً. والمطلوب تبويب البيانات في جدول توزيع تكراري بسيط .

41، 40، 91، 75، 76، 64، 60، 42، 56، 52، 50، 98، 61، 68، 63، 65، 68، 70، 60، 83، 84، 88، 51، 73، 75، 79، 80، 67، 63، 66، 51

الحل (2.3):

- نحدد عدد الفئات باستخدام الصيغة (1.3) $K=1+3.322 (\log n)$ يكون لدينا :

$$K=1+3.322 (\log 31) \\ \approx 6$$

- ايجاد طول الفئة التكرارية $H = \frac{\text{اكبر قيمة (98) - أصغر قيمة (40)}}{\text{عدد الفئات (6)}} \approx 10$

- تعيين حدود الفئات: من البيانات اعلاه نجد اقل قيمة هي 40 فتكون هي الحد الأدنى للفئة الأولى، اما الحد الأعلى فهو حسيلة جمع قيمة الحد الأدنى الى طول الفئة 10 مطروحا منه 1 ليصبح مقداره 49، اما الحد الأدنى للفئة الثانية فهي القيمة اللاحقة للحد الأعلى للفئة السابقة وهي 50، وبإضافة طول الفئة مطروحا منه 1 الى الحد الأدنى لذات الفئة لنحصل على قيمة الحد الأعلى للفئة الثانية وهي 59 وهكذا مع باقي الفئات، فنحصل على الفئات المبينة في جدول رقم (14.3) .

- توزيع التكرارات على الفئات: وفيها يتم تأشير كل تكرار مقابل الفئة المناسبة له، ومن ثم القيام بجمع هذه الاشارات ووضعها في حقل التكرار كما هو مبين في الجدول (14.3) ادناه:

جدول رقم (14.3)

جدول توزيع تكراري

| التكرار (fi) | الاشارات | الفئات |
|------------------|---------------|---------|
| 3 | 111 | 49 -40 |
| 5 | 11111 | 59 -50 |
| 11 | 1 11111 11111 | 69 -60 |
| 6 | 1 11111 | 79 -70 |
| 4 | 1111 | 89 -80 |
| 2 | 11 | 99 -90 |
| $\Sigma fi = 31$ | | المجموع |

الفئات المفتوحة والفئات غير المتساوية الأطوال

في حالات معينة يصادف ان تضم مجموعة البيانات بعض القيم المتطرفة او المتباينة مع اتجاه القيم الاخرى، فاذا كانت متطرفة في الصغر فستخصص الفئة الاولى، وعندما تكون متطرفة في الكبر فسيعلق الامر بآخر فئة، مما يستوجب إما جعل بعض الفئات غير متساوية الطول، أو القيام بشمول فئات اضافية، فالفئات غير المتساوية الطول تخلق صعوبة في اعطاء صورة واضحة عن شكل التوزيع عند اجراء المقارنة الفئوية. وفي حالة جعل كافة الفئات متساوية الطول سيؤدي الامر الى أن تكون بعض الفئات خالية من التكرار. ومن الخيارات الممكنة لمعالجة الحالة الاولى هو رفع الحد الادنى من الفئة الاولى اذا كان التطرف في الصغر، ورفع الحد الاعلى من الفئة الاخيرة اذا كان التطرف في الكبر، وهو ما يدعى بالفئات المفتوحة كما هو مبين في الجدول رقم (15.3):

جدول رقم (15.3)
فئات مفتوحة

| التكرار fi | الفئات |
|------------|-----------|
| 4 | 160 فأقل |
| 7 | 166 - 161 |
| 12 | 172 - 167 |
| 6 | 178 - 173 |
| 5 | 184 - 197 |
| 2 | 185 فأكثر |

ولمعالجة الحالة الثانية (الفئات غير المتساوية الطول)، يتعين علينا تعديل التكرارات قبل البدء بحساب المقياس ويتم ذلك باستخدام طريقة شبرد Sheppard's correction of grouping وذلك بقسمة التكرار الخاص بكل فئة على طول الفئة المقابلة له للحصول تكرارات جديدة يتم اعتمادها في حساب المقياس المطلوب .

الحدود الحقيقية للفئات (نهايات الفئة)

رغم ان الفئات تضم كافة البيانات عند تبويبها الا انها غير متصلة ببعضها، مما يجعلها متغيرا متقطعا Discrete variable، اي ان هناك مديات فاصلة بين فئة واخرى، مما له تاثير مباشر في التوزيعات الاحتمالية عند تمهيد المنحنى التكراري. ولأجل تعديل حدي الفئات يتم اعادة حساب حدي الفئة للحصول على ما يسمى بنهايات الفئات كالآتي :

$$\frac{\text{الحد الأعلى للفئة السابقة} + \text{الحد الأدنى للفئة المعنية}}{2} = \text{النهاية الدنيا للفئة المعنية}$$

$$\frac{\text{الحد الأعلى للفئة المعنية} + \text{الحد الأدنى للفئة اللاحقة}}{2} = \text{النهاية العليا للفئة المعنية}$$

فمثلا الحدود الحقيقية لفئات الجدول رقم (14.3) تصبح كما في الجدول (16.3) ادناه :

جدول رقم (16.3)

| الفئات | الفئات الحقيقية | النهاية الدنيا | النهاية العليا |
|---------|-----------------|----------------|----------------|
| 49 - 40 | 49.5 - 39.5 | 39.5 | 49.5 |
| 59 - 50 | 59.5 - 49.5 | 49.5 | 59.5 |
| 69 - 60 | 69.5 - 59.5 | 59.5 | 69.5 |
| 79 - 70 | 79.5 - 69.5 | 69.5 | 79.5 |
| 89 - 80 | 89.5 - 79.5 | 79.5 | 89.5 |
| 99 - 90 | 99.5 - 89.5 | 89.5 | 99.5 |

2. التوزيع التكراري المتجمع Cumulative Frequency Distribution

وتعود أهمية التكرار المتجمع عندما ينصب الاهتمام على العدد الذي يزيد أو يقل عن قيمة معينة، فمثلا قد يهمنا من توزيع علامات الطلبة لمعرفة الذين تقل علاماتهم عن 50 أو أولئك الذين تزيد على 80 وهكذا. وهناك نوعان من التكرارات المتجمعة، فتلک التي يبدأ تجميعها من الأعلى باتجاه الأسفل ويصطلح على تسميتها بالمتجمع الصاعد، وفيه نبدأ بأول تكرار وعند الثاني نضيف إليه التكرار البسيط الثاني وفي الثالث نضيف للمتجمع الثاني التكرار البسيط الثالث وهكذا. أما النوع الآخر وهو المتجمع النازل فيكون آخر تكرار هو الأول ثم نضيف إليه التكرار البسيط قبل الأخير ليصبح التكرار المتجمع الثاني من الأخير وهكذا، ليكون عند أعلى فئة مساويا لمجموع التكرار البسيط. وبالرجوع إلى جدول التوزيع التكراري البسيط رقم (16-3) يكون لدينا التوزيعات المتجمعة الصاعدة والنازلة كما هو مبين في الجدول رقم (17-3):

جدول رقم (173)
التوزيع التكراري التجميع

| المتجمع النازل | المتجمع الصاعد | التكرار البسيط | الفئات |
|----------------|----------------|----------------|---------|
| 31 | 3 | 3 | 49 - 40 |
| 28 | 8 | 5 | 59 - 50 |
| 23 | 19 | 11 | 69 - 60 |
| 12 | 25 | 6 | 79 - 70 |
| 6 | 29 | 4 | 89 - 80 |
| 2 | 31 | 2 | 99 - 90 |

وحيث ان القراءة الصحيحة للتكرارات المتجمعة تقتزن بالحدود الحقيقية للفئات، وذلك باستخدام الحدود الحقيقية العليا مع المتجمع الصاعد والحدود الحقيقية الدنيا مع المتجمع النازل كما هو مبين في الجدولين رقم (3-18) و (3-19) التالية:

جدول رقم (183)
قراءة التجمع الصاعد جدول

| التكرار الصاعد | الحدود الحقيقية العليا |
|----------------|------------------------|
| 0 | أقل من 39.5 |
| 3 | أقل من 49.5 |
| 8 | أقل من 59.5 |
| 19 | أقل من 69.5 |
| 25 | أقل من 79.5 |
| 29 | أقل من 89.5 |
| 31 | أقل من 99.5 |

جدول رقم (3-19)
قراءة التجميع النازل

| الحدود الحقيقية الدنيا | التكرار النازل |
|------------------------|----------------|
| 39.5 فأكثر | 31 |
| 49.5 فأكثر | 28 |
| 59.5 فأكثر | 23 |
| 69.5 فأكثر | 12 |
| 79.5 فأكثر | 6 |
| 89.5 فأكثر | 2 |
| 99.5 فأكثر | 0 |

التكرار النسبي البسيط والمتجمع

فالتكرار النسبي البسيط و المتجمع هو عبارة عن نسبة ما يشكله تكرار كل فئة من المجموع. وتأتي أهمية النسب عندما تكون قيم التكرارات كبيرة جدا فتصبح النسب اسهل في المقارنة بدلا من الارقام، وتحقق عملية التحويل الى نسب من خلال قسمة تكرار كل فئة على مجموع التكرارات وضربها بـ 100.

3 التوزيع التكراري المزدوج Paired Frequency Distribution

- ويستعمل هذا النوع من التوزيع في تبويب البيانات في حالة وجود ظاهرتين (متغيرين) تعتمد كل منهما على الأخرى كاطوال الاشخاص واوزانهم او كمية بضاعة ما وسعرها وما شابه. ويتم بناء هذه الجداول حسب الخطوات التالية:
- تحديد عدد واطوال فئات كل من المتغيرين بصورة مستقلة باستخدام نفس الاجراءات السابقة المتعلقة بالتوزيع التكراري البسيط.
 - ترتيب فئات احد المتغيرين افقيا والآخر عموديا في الجدول .
 - تبويب البيانات على الفئات، بوضع الرقم في الخانة التي تعود لفئتي المتغيرين ذات العلاقة بذلك الرقم.

- يخصص حقلان في نهاية الجدول أحدهما أفقي لمجاميع المتغير الأول، والآخر عمودي لمجاميع المتغير الثاني، وذلك بغية التأكد من مساواة كلا المجموعين.

مثال (2-4): البيانات التالية تمثل علامات الطلبة البالغ عددهم 31 طالباً في مادة الاحصاء، ومعدل كل منهم في الثانوية العامة، والمطلوب تبويب البيانات في جدول توزيع تكراري مزدوج باستخدام 6 فئات لمتغير علامات مادة الاحصاء و 5 فئات لمعدلات الثانوية العامة.

علامات مادة الاحصاء: 41، 40، 91، 75، 76، 64، 60، 43، 56، 52، 50، 95، 61، 68، 63، 65، 68، 70، 60، 83، 84، 88، 51، 73، 75، 79، 80، 67، 63، 66، 51

معدلات الثانوية العامة: 61، 70، 71، 69، 65، 59، 58، 56، 60، 65، 68، 78، 72، 65، 59، 62، 60، 72، 70، 80، 87، 83، 54، 58، 61، 67، 69، 60، 58، 62، 57

الحل: (2-4)

| فئات معدل الثانوية العامة | | | | | | فئات علامات |
|---------------------------|---------|---------|---------|---------|---------|-------------|
| المجموع | 83 - 78 | 77 - 72 | 71 - 66 | 65 - 60 | 59 - 54 | الاحصاء |
| 3 | | | 1 | 1 | 1 | 49 - 40 |
| 5 | | | 1 | 11 | 11 | 59 - 50 |
| 11 | | 1 | 1 | 11111 | 1111 | 69 - 60 |
| 6 | | 1 | 1 | 111 | 1 | 79 - 70 |
| 4 | 111 | | 1 | | | 89 - 80 |
| 2 | 1 | | 1 | | | 99 - 90 |
| 31 | 4 | 2 | 6 | 11 | 8 | المجموع |

4. التوزيعات النوعية (الوصفية) والزمنية والجغرافية

ولا تحتاج هذه التوزيعات الى فئات، بل ان توزيعها يكون حسب الصفة التي تعود اليها البيانات كالمهنة او الحالة التعليمية مثلاً، او وحدة زمنية كالاشهر، او وحدة جغرافية كالمدن او الاقاليم. وتتطلب هذه الجداول مراعاة شروط او مواصفات معينة اهمها:

- ترقيم الجدول، وان يكون الترقيم مشتقاً من الفصل او الباب الذي يعود اليه.
- عنوان للجدول يدل على محتوياته وعلى طبيعة تصنيفاته والوحدة القياسية المستخدمة في قياس بياناته والزمن والمكان الذي يعود اليه.
- هوامش سفلية اذا اقتضى شرح احد او بعض بياناته.
- مصدر البيانات لتسهيل الرجوع اليها عند الحاجة او للاطمئنان لدقة البيانات.

ومن الامثلة على هذه الانواع من الجداول توزيع السكان حسب المحافظات او توزيع عدد حوادث الطرق حسب نوع الحادث او حسب نوع واسطة النقل، او تطور الاستيرادات والصادرات حسب السنين او عدد الطلبة حسب الجامعات وغيرها .

5. العرض البياني

وحيث قد تم عرض اهم الاشكال البيانية في فقرة استخدام الحاسوب في بداية هذا الفصل، فسيتم في هذه الفقرة تناول الاجراءات المطلوبة لإعداد هذه الاشكال البيانية في حالة القيام بانجازها يدوياً .

اولاً: الاشكال البيانية للبيانات المبوبة على شكل فئات

1- المضلع والمنحنى التكراري:

تجدر الإشارة الى ان مساحة ما يسمى بالمدرج التكراري Histogram هي ذات مساحة المضلع التكراري، لذا فان الاستعانة بعرض المضلع او المنحنى التكراري (الذي هو تمهيد للمضلعات) هو الحصيلة النهائية التي يتم الركون اليها للمدراج التكراري خاصة اذا ما علمنا بان مساحة المضلع او المنحنى هي الهدف النهائي من المدرج. إن رسم المضلع يتم بتحديد مراكز الفئات (الحد الأدنى + الحد الأعلى للفئة

مقسومة على 2) على المحور الافقي، وتعيين التكرارات على المحور العمودي ومن ثم التوصيل بين نهايات النقاط التي يتم تحديدها بخطوط مستقيمة، وفي حالة تمهيد نقاط التقاء المستقيمات نحصل على المنحنى التكراري كما هو مبين في الاشكال البيانية (19.3) و(20.3) . ومن خصائص المضلع او المنحنى التكراري امكانية رسم اكثر من مضلع او منحنى في نفس الشكل البياني .

2- المضلع والمنحنى التكراري المتجمع :

ان رسم المضلع المتجمع التكراري يتم بتثبيت قيم المتجمع الصاعد او النازل على المحور العمودي، والنهايات العليا للفئات الحقيقية او مراكز الفئات على المحور الافقي، ومن ثم توصيل خطوط مستقيمة بين النقاط التي يتم تعيينها، وبتمهيد نقاط التقاء المستقيمات نحصل على المنحنى التكراري المتجمع، والشكل البياني رقم (3-21) يوضح الشكل الذي يظهر عليه المضلع التكراري المتجمع.

ثانيا: الاشكال البيانية للبيانات غير المبوبة

1- الاعمدة والمستطيلات البيانية

وهي من اكثر الاشكال البيانية استخداما وتخص البيانات التي تكون مشاهداتها بصيغة صفات او وحدات زمنية، كالسنين والاشهر والايام او جغرافية كالمدن والاقاليم والدول. ويتم رسمها بتثبيت السنين او الصفات او غيرها على المحور الافقي، والتكرارات على المحور العمودي، وبذلك فان اطوال الاعمدة الناتجة تمثل العلاقة بين كل صفة او سنة او مدينة وتكرارها. والاعمدة على عدة انواع منها الاحادية (البسيطة) وتخص متغيراً واحداً كما هو مبين في الشكل رقم (3-22)، وقد تكون من نوع الاعمدة المتعددة وتستخدم لعرض متغيرين (ظاهرتين) او اكثر كما يوضحه الشكل البياني رقم (3-23). وعندما تعرض عدة ظواهر (متغيرات) او عدة مستويات للظاهرة الواحدة في ذات العمود، يطلق عليها الاعمدة المركبة، بحيث يمثل ارتفاع العمود مجموع قيم الظواهر او مجموع مستويات الظاهرة الواحدة، كما هو موضح في الشكل البياني رقم (3-24) .

2- الدائرة البيانية

وتستخدم عندما يكون الهدف إبراز الاجزاء التي تتكون منها الظاهرة، الا انها لا تستخدم اذا كان الهدف متابعة تطور التغييرات التي تطرا على الظاهرة. وانجازها يتم بتقسيم مساحة الدائرة الى قطاعات، كل قطاع يمثل جزءا او احد مكونات الظاهرة. ويتم تحديد كل جزء من خلال ضرب الزاوية المركبة للدائرة والتي مقدارها 360 بحاصل قسمة الجزء المعني على مجموع قيم الاجزاء، اي :

$$\text{مساحة الجزء المعني} = \frac{\text{قيمة الجزء}}{\text{مجموع قيم الأجزاء}} \times 360$$

كما هو مبين في الشكل البياني (3-25).

3- الرسوم والصور

يعتمد اعداد الرسوم والصور على شكل وحدات الظاهرة المعنية بالدراسة كاساس في اختيار الرسم او الصورة، وافترض قيمة محددة لكل وحدة من وحدات الظاهرة. فمثلا اذا كنا بصدد عرض تطور عدد السيارات، فسنختار صورة السيارة كمقياس للتعبير، واذا كنا بصدد عرض عدد السكان فنختار صورة تخطيطية لشخص، وللتعبير عن عدد المساكن يتم اعتماد صورة رمزية لمسكن وهكذا. والشكل البياني رقم (3-26) يمثل نموذجا للرسوم والصور باستخدام الامر الفرعي symbol، والامر الفرعي clip art من الامر الرئيسي Insert.

تمارين الفصل الثالث

تمرين (1.3): قام احد مصانع المواد الغذائية المعلبة باخذ عينة من الانتاج لأحد انواع منتجاته بهدف التأكد من تحقق الوزن المقرر البالغ 50 غم للعلبة الواحدة، وكان حجم العينة 100 علية، وبعد اجراء عملية الوزن كانت النتائج مبين في ادناه :

36 ، 37 ، 38 ، 38 ، 35 ، 39 ، 39 ، 36 ، 40 ، 41 ، 35 ، 40 ، 41 ، 45 ،
41 ، 41 ، 42 ، 50 ، 52 ، 53 ، 42 ، 42 ، 42 ، 42 ، 42 ، 43 ، 43 ، 43 ،
43 ، 43 ، 47 ، 44 ، 44 ، 44 ، 44 ، 44 ، 46 ، 43 ، 44 ، 44 ، 45 ، 45 ،
45 ، 45 ، 46 ، 45 ، 46 ، 48 ، 51 ، 46 ، 46 ، 47 ، 47 ، 47 ، 48 ،
48 ، 48 ، 48 ، 48 ، 47 ، 49 ، 49 ، 49 ، 49 ، 50 ، 49 ، 50 ، 50 ،
49 ، 52 ، 51 ، 52 ، 51 ، 53 ، 53 ، 53 ، 55 ، 48 ، 49 ، 51 ،
51 ، 48 ، 46 ، 54 ، 50 ، 53 ، 53 ، 44 ، 45 ، 45 ، 45 ، 46 ،
45،55

والمطلوب :

1. تبويب البيانات باستخدام برنامج SPSS.
2. تبويب البيانات في جدول توزيع تكراري عدد فئاته 7 باستخدام برنامج EXCEL.
3. استخدام برنامج EXCEL لعرض بيانات التوزيع التكراري موضوع (2) اعلاه على شكل مضلع تكراري، وبرنامج SPSS لعرض المدرج التكراري.
4. ايجاد التكرار المتجمع الصاعد والنازل يدويا.
5. عرض بيانات التكرار المتجمع الصاعد والنازل على شكل منحنيات تكرارية باستخدام برنامج EXCEL.
6. ايجاد مراكز الفئات والحدود الحقيقية للفئات.

تمرين (2.3): اعرض بيانات الجدول التالي الذي يضم عدد وسائل النقل الافتراضية للفترة 1997-2002 في الاشكال البيانية التالية مستخدماً برنامج EXCEL ويدوياً:

1. الاعمدة البيانية المتعددة والمركبة

2. الدائرة البيانية لسنة 2002

| السنة | بيك اب | صالون | لوري |
|-------|--------|-------|------|
| 1997 | 5575 | 6494 | 1848 |
| 1998 | 5764 | 6757 | 2056 |
| 1999 | 7338 | 12446 | 2216 |
| 2000 | 9600 | 17073 | 2312 |
| 2001 | 10311 | 16466 | 2508 |
| 2002 | 9191 | 16899 | 2596 |



مقاييس النزعة المركزية (المتوسطات) والتشتت

Central Tendency Measures & Desperation

4.1 مقدمة Introduction

هناك خاصيتان أساسيتان لأية بيانات احصائية تساعد على إعطاء مدلول واضح لوصفها هما: النزعة المركزية ومقاييسها متمثلة بالمتوسطات التي بواسطتها نتمكن من تحديد موقع النقطة التي تتمحور حولها كثافة القيم. أما الثانية فهي مقاييس التشتت التي يقصد بها حالة الانتشار التي تكون عليها البيانات حول المركز (المتوسط). والمتوسط هو قيمة مفردة تمثل مجموعة من قيم المعطيات، وهناك عدة أنواع من المتوسطات لكل منها طريقته الخاصة في الاحتساب والتي تم تناولها عند التطرق الى الطريقة اليدوية لاحقا من هذا الفصل، وهذه الأنواع هي: الوسط الحسابي Arithmetic mean – الوسيط Median – الوسط الهندسي Geometric mean – المنوال Mode – الوسط التوافقي Harmonic mean، الا ان المتوسط الأخير قليل الاستخدام.

4.2 استخدام الحاسوب مع برنامج SPSS

إن الحصول على مقاييس كل من النزعة المركزية (المتوسطات) والتشتت باستخدام برنامج SPSS يمكن انجازه من خلال الامر الرئيسي Analyze وعبر اكثر من امر فرعي منه، كما اتضح لنا في الفصل الثالث عند اجراء عملية تبويب البيانات، ويتم ذلك بتوظيف الامر الفرعي Reports ومن ثم اختيار الطريقة Case summaries الذي حصلنا بواسطتها على كافة مقاييس النزعة المركزية والتشتت. اما الامر الفرعي الآخر فهو Descriptive Statistics ثم اختيار الطريقة Frequencies ومنه نحصل على كافة التفاصيل المتعلقة بانواع المتوسطات والتشتت كما هو مبين في مخرجات الجدول رقم (2-3). وانجاز ذلك وكما اسلفنا في الفصل الثالث يتم باخضاع جدول المدخلات رقم (2.3) و استدعاء الامر الرئيسي Analyze ومنه نختار الامر الفرعي Descriptive Statistics وبعد ذلك التأشير على طريقة Frequencies.

3.4 الطريقة اليدوية

ان استخراج قيمة المتوسطات يكون إما من قيم غير مبوبة (ungrouped data) أي بشكل وحدات لكل منها قيمتها الخاصة، أو من قيم مبوبة (grouped data) وتكون على شكل جداول تكرارية، وكل تكرار يمثل عدد الوحدات التي تقع ضمن فئة معينة لها نهايتان دنيا وعليا، دون تحديد القيم الفعلية لتلك الوحدات.

1- الوسط الحسابي The Arithmetic Mean :

1- حالة المعطيات غير المبوبة Ungrouped Data :

إذا كان لدينا مجموعة قيم هي X_1, X_2, \dots, X_n فإن وسطها الحسابي ولنرمز له \bar{X} في حالة العينة n ، ونرمز له μ في حالة المجتمع N ، سيكون عبارة عن مجموع هذه القيم مقسومة على عددها، أي في حالة العينة

$$\bar{X} = \frac{\sum_{i=1}^n xi}{n} \quad (1.4)$$

حيث إن:

$\sum xi$ هي مجموع قيم المفردات و n هي عدد المفردات (حجم العينة)

وفي حالة المجتمع فإن صيغة الوسط الحسابي μ هي :

$$\mu = \frac{\sum_{i=1}^N xi}{N} \quad (2.4)$$

مثال (1.4): أوجد الوسط الحسابي لعدد العاملين في 5 مخازن مختلفة، إذ كان عددهم في هذه المخازن هو على التوالي 3، 5، 6، 4، 6:

الحل (1.4):

$$\sum xi = 3 + 5 + 6 + 4 + 6 \\ = 24$$

وبتطبيق الصيغة (1.4) نحصل على :

$$\bar{X} = \frac{\sum xi}{n}$$

$$= \frac{24}{5} = 4.8$$

وهو قيمة الوسط الحسابي لعدد العاملين أي متوسط عدد العاملين في كل مخزن.

ولكن عندما تكون القيم غير متساوية من حيث أهميتها، عندها يتطلب ترجيح القيم بما يتناسب وأهمية كل منها، وتصبح صيغة احتساب الوسط الحسابي المرجح (أوالموزون) ولنرمز له بـ \bar{X}_w كالآتي :

$$\bar{X}_w = \frac{\sum x_i w_i}{\sum w_i}$$

حيث إن:

w_i هو وزن القيمة x_i

مثال (2.4): باع أحد اصحاب محلات الفاكهة نوعاً من الفاكهة بثلاثة اسعار مختلفة كما في ادناه:

| السعر (دينار/ كيلو) | الكمية المباعة (كيلو) |
|---------------------|-----------------------|
| 3.00 | 5 |
| 0.60 | 60 |
| 0.20 | 300 |

والمطلوب إيجاد الوسط الحسابي المرجح لسعر البيع :

الحل (2.4) :

حيث أن الكمية هي أوزان لترجيح الأسعار، يكون لدينا :

$$x_1 = 3.0, x_2 = 0.60, x_3 = 0.20$$

$$w_1 = 5, w_2 = 60, w_3 = 300$$

وبتطبيق الصيغة (3.4) نحصل على:

$$\begin{aligned} \bar{X}_w &= \frac{\sum x_i w_i}{\sum w_i} \\ &= \frac{(3)(5) + (0.60)(60) + (0.20)(300)}{5 + 60 + 300} = 0.304 \end{aligned}$$

2 - حالة المعطيات المبوبة Grouped Data

أما عند التعامل مع بيانات مبوبة تعود لفئات لها مدى يقع بين حد أدنى وحد أعلى، فسنفترض بأن المعطيات تقع في مركز الفئة، أي أن تكرار كل فئة سيقع قسم منه تحت نقطة مركز الفئة، والقسم الآخر فوق نقطة المركز، وبذلك فستكون في المعدل عند نقطة مركز الفئة، فإذا رمزنا لمراكز الفئات بـ x_i فستكون عبارة عن حاصل قسمة مجموع حدي الفئة على 2 والوسط الحسابي يتم استخراجه باستخدام الصيغة التالية :

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i} \quad (4.4)$$

مثال (3.4) : أوجد الوسط الحسابي لمعطيات جدول التوزيع التكراري التالي :

| الفئات | التكرار (f_i) |
|--------|-------------------|
| 40-49 | 3 |
| 50-59 | 5 |
| 60-69 | 11 |
| 70-79 | 6 |
| 80-89 | 4 |
| 90-99 | 2 |
| | $\sum f_i = 31$ |

الحل (3.4) :

أ- نستخرج مراكز الفئات x_i ونحتسب القيم $\sum x_i f_i$ فيكون لدينا :

| x_i | f_i | $x_i f_i$ |
|-------|-----------------|-------------------------|
| 44.5 | 3 | 133.5 |
| 54.5 | 5 | 272.5 |
| 64.5 | 11 | 709.5 |
| 74.5 | 6 | 447 |
| 84.5 | 4 | 338 |
| 94.5 | 2 | 189 |
| | $\sum f_i = 31$ | $\sum x_i f_i = 2089.5$ |

ب- نطبق الصيغة (4.4) نحصل على:

$$\bar{X} = \frac{\sum x_i f_i}{\sum f_i}$$

$$= \frac{2089.5}{31} = 67.4$$

أما في حالة أن تتضمن الجداول التكرارية لفئات مفتوحة، فيمكن افتراض بأنها ستأخذ نفس طول (أو المدى) الفئة المعتمدة مع الفئات المتبقية، ونحدد قيمة حد الفئة المفتوحة، ومن ثم نطبق نفس إجراءات الاحتساب للمثال أعلاه، وكما مبين في المثال (4.4) أدناه :

مثال (4.4): الجدول التالي يتضمن الأجر الشهري بالدينار لـ 282 عاملاً في إحدى الشركات الصناعية. والمطلوب احتساب الوسط الحسابي لأجر العامل الواحد شهرياً

| فئات الأجور | عدد العمال f_i |
|-------------|------------------|
| 109-80 | 26 |
| 139-110 | 78 |
| 169-140 | 122 |
| 199-170 | 34 |
| 229-200 | 14 |
| 230 فأكثر | 8 |
| المجموع | $\sum f_i = 282$ |

الحل (4.4) :

- أ. نفترض بأن الفئة الأخيرة تنتهي بالأجر 259 دينار وذلك بالاستناد إلى طول الفئة المعتمد لباقي الفئات وهو 30.
- ب. نستخرج مراكز الفئات x_i ونحتسب قيمة $\sum x_i f_i$ وكالاتي:

| Xi fi | Fi (التكرار) | Xi (مراكز الفئات) |
|------------------------------|-----------------|-------------------|
| 2457 | 26 | 94.5 |
| 9711 | 78 | 124.5 |
| 18849 | 122 | 154.5 |
| 6273 | 34 | 184.5 |
| 3002 | 14 | 214.5 |
| 1956 | 8 | 244.5 |
| $\sum xi \text{ fi} = 42249$ | $\sum fi = 282$ | |

ج. نطبق الصيغة (4.4) فنحصل على :

$$\bar{X} = \frac{\sum xi \text{ fi}}{\sum fi}$$

$$= \frac{42249}{282} = 149.82$$

وهو متوسط الاجر الشهري بالدينار

وعندما نواجه جداول تكرارية مطولة أو معقدة، فبالإمكان تطبيق طريقة مختصرة لاحتساب الوسط الحسابي، وذلك باستخدام قيمة أصل اعتباطية، تدعى بالقيمة الفرضية ونرمز لها x_0 . فمثلا بدلا من اعتماد القيم:

94.5، 124.5، ... الخ كمراكز فئات بالنسبة للجدول التكراري موضوع المثال (4.4)، بالإمكان أخذ القيمة 154.5 واعتبارها قيمة اصل فرضية، وتدوين صفر بدلا من عنها. حيث أن قيم مراكز الفئات الباقية هي اما اقل أو اعلى من 154.5 وجميعها بطول فئة مقدارة 30، وبذلك سنقل الانحراف الى مرتبة واحدة. ويمكن تلخيص الطريقة بالآتي :

اولا: نحدد القيمة الفرضية x_0 كنقطة اصل، ويتم ذلك اعتباطيا.

ثانيا: نحسب الانحراف D_i لـ x_i عن القيمة الفرضية X_0 مقسومة على طول الفئة، H أي:

$$D_i = \frac{X_i - X_o}{H}$$

ثالثاً: نستخرج $\sum D_i f_i$ أي مجموع حاصل ضرب التكرار f_i بالانحراف D_i .
رابعاً: نحتسب قيمة الوسط الحسابي المفترض \bar{X}_o (fictitious Mean) كالآتي:

$$\bar{X}_o = \frac{\sum D_i f_i}{\sum f_i}$$

خامساً: نحول الوسط الحسابي المفترض \bar{X}_o الى الوسط الحسابي الحقيقي \bar{X} وذلك:

$$\bar{X} = X_o + \bar{X}_o H \quad (7.4)$$

وباستخدام الطريقة المختصرة بموجب الخطوات اعلاه، مع المثال (4.4) يكون

لدينا :

| الفئات | مراكز الفئات X_i | الانحراف عن القيمة الفرضية $D_i = \frac{X_i - X_o}{H}$ | التكرار f_i | $D_i f_i$ |
|--------|-----------------------|--|------------------|--------------------|
| 40-49 | 44.5 | -2 | 3 | -6 |
| 50-59 | 54.5 | -1 | 5 | -5 |
| 60-69 | 64.5 | 0 | 11 | 0 |
| 70-79 | 74.5 | 1 | 6 | 6 |
| 80-89 | 84.5 | 2 | 4 | 8 |
| 90-99 | 94.5 | 3 | 2 | 6 |
| | | | $\sum f_i = 31$ | $\sum D_i f_i = 9$ |

ب - نستخرج الوسط الحسابي الفرضي باستخدام الصيغة :

$$\bar{X}_o = \frac{\sum D_i f_i}{\sum f_i}$$

$$= \frac{9}{31} = 0.29$$

ج - نحتسب الوسط الحسابي الحقيقي باستخدام الصيغة (7.4):

$$\begin{aligned}\bar{X} &= X_o + \bar{X}_o H \\ &= 64.4 + (0.29)(10) = 67.4\end{aligned}$$

وتجدر الإشارة الى أن الطريقة المختصرة لا يمكن استخدامها مع التوزيعات التكرارية غير المتساوية في اطوال فئاتها، إلا بعد اجراء التعديلات التي اشرنا إليها في الفصل الثالث والمتمثلة بقسمة تكرار كل فئة على طولها للحصول على تكرارات جديدة.

3- خواص الوسط الحسابي وعيوبه Arithmetic Mean Properties :

اولا: العمليات الحسابية المطلوبة لاحتسابه غير معقدة، رغم انها طويلة نسبيا مقارنة بالعمليات المطلوبة للمتوسطات الأخرى.

ثانيا: عملية احتسابه مفهومة لسعة استخداماته.

ثالثا: عملية احتسابه تشتمل على كافة وحدات التوزيع التكراري.

رابعا: امكانية توظيفه لاجاد مجموع قيم المشاهدات، $\sum X_i$ عند معلومية حجم العينة n حيث إن :

$$\begin{aligned}\bar{X} &= \frac{\sum X}{n} & \text{or} & & \mu &= \frac{\sum X}{N} \\ \sum X &= n\bar{X} & & & \sum X &= N\mu\end{aligned}$$

فمثلا اذا كان عدد زبائن أحد المخازن هو $N=300$ وان متوسط مشتريات الزبون الواحد هو $u=58$ دينار فإن مجموع مبيعات المخزن هي :

$$\begin{aligned}\sum X &= Nu = (300)(58) \\ &= 17400 \text{ دينار}\end{aligned}$$

وبصورة عامة فإن الوسط الحسابي يعتبر أفضل إحصاء لتمثيل النزعة المركزية، لاسأسه النظري الذي يسمح لاستخدامه في التحليلات الاحصائية المتقدمة، فلانحراف عن الوسط الحسابي ميزتان على غاية الأهمية، هي أن مجموع هذه الانحرافات تؤول الى الصفر وأن مجموع مربعات هذه الانحرافات هي أقل ما يمكن. كما أن الانحرافات عن الوسط الحسابي تجهز معلومات اساسية لأي توزيع احتمالي، فاذا ما سحبنا العينات من مجتمع ماء، نجد أن الوسط الحسابي هو أقل تذبذباً عما هو عليه مع مقاييس النزعة المركزية الأخرى، فهو بذلك أفضل تقدير لمعلمة المجتمع.

أما ابرز عيوبه فيمكن اجمالها بـ :

أولاً: قابليته للتأثر بعدد قليل من الوحدات المتطرفة وبالتالي يصبح غير ممثل للبيانات.

ثانياً: لا يمكن قياسة والتأكد منه بالطرق البيانية.

2. الوسيط The Median

1- حالة المعطيات غير المبوبة، Ungrouped Data :

عندما يتم ترتيب المعطيات تصاعدياً من الأصغر فالأكبر، أو تنازلياً من الأكبر فالأصغر، فإن الوسيط يصبح عبارة عن القيمة الوسطية (عندما يكون عدد المعطيات فردياً)، أو قيمة متوسط القيمتين الوسطيتين (عندما يكون عدد المعطيات زوجياً). وبذلك فإن تحديد قيمة الوسيط ولنرمز له بـ M_d يتم كالآتي :

أولاً: عندما يكون العدد فردياً، فإن موقع قيمة الوسيط يكون في الترتيب :

$$\frac{n+1}{2} \quad (8.4)$$

ثانياً: أما عندما يكون اللعدد زوجياً، فإن موقع القيمة الأولى للوسيط هو في

الترتيب :

$$\frac{n}{2} \quad (9.4)$$

وموقع القيمة الثانية للوسيط هو في الترتيب :

$$\frac{n+2}{2} \quad (10.4)$$

وبذلك تكون قيمة عبارة عن متوسط القيمتين (الأولى والثانية).

مثال (5.4): في خمس اختبارات في مادة الإحصاء حصل أحد طلبة الكيمياء على النتائج التالية 91,80,86,75,94. اوجد الوسيط لهذه الدرجات.

الحل (5.4):

أ- نرتب المعطيات تصاعديا فيكون لدينا: 75,80,86,91,94

ب- نحدد موقع الوسيط، وحيث إن عدد المعطيات فرديا نستخدم الصيغة (8.4)، وأن القيمة الواقعة في ذلك الموقع تمثل الوسيط.

$$\frac{n+1}{2} = \frac{6}{2} = 3$$

وعليه فإن القيمة الواقعة في الترتيب الثالث وهي 86 تمثل الوسيط.

مثال (6.4): عند فحص النيكوتين لعينة من أحد انواع السكاثر، وجد ان كميتها (بالمغم) هي 2.1,3.2,2.9,2.6,2.8,2.4 فما هو الوسيط.

الحل (6.4):

أ- نرتب المعطيات تصاعديا فيكون لدينا: 2.1,2.4,2.4,2.8,2.9,3.2

ب- نحدد موقع الوسيط، وحيث إن عدد المعطيات زوجيا نستخدم الصيغتين (9.4) و(10.4) لذلك، فإن متوسط القيمتين الواقعتين في المواقع المستخرجة تمثل الوسيط وكالاتي:

$$\frac{n}{2} = \frac{6}{2} = 3 \quad \text{موقع القيمة الأولى:}$$

$$\frac{n}{2} = \frac{6}{2} = 3 \quad \text{موقع القيمة الثانية:}$$

نحتسب متوسط القيمتين الواقعتين في الترتيب الثالث والترتيب الرابع وهي 2.4, 2.8 على التوالي فنحصل على قيمة الوسيط :

$$Md = \frac{2.4 + 2.8}{2} = 2.6$$

2- حالة المعطيات المبوبة: Grouped data

أولاً: الطريقة الحسابية

- أ. نستخرج التوزيع التكراري المتجمع الصاعد.
- ب. نحدد موقع الوسيط بقسمة مجموع التكرارات على 2، أي $\frac{\sum f_i}{2}$
- ج. نحدد قيمة موقع التكرار الوسيط بين التكرارات المتجمعة.
- د. نحدد الفئة الوسيطة، فإذا كانت قيمة موقع الوسيط مساوية لأي تكرار متجمع حينئذ فإن فئة ذلك التكرار ستكون هي الفئة الوسيطة، أما إذا وقعت بين تكرارين متجمعين فإن الفئة اللاحقة لقيمة الموقع ستكون هي الفئة الوسيطة.
- هـ. نستخدم الصيغة التالية لاحتساب قيمة الوسيط :

$$Md = L + \frac{\frac{\sum f_i}{2} - f_l}{f_2 - f_l} H \quad (11.4)$$

حيث إن :

L: الحد الأدنى لفئة الوسيط.

$\frac{\sum f_i}{2}$: قيمة موقع الوسيط.

f_l: التكرار المتجمع السابق لقيمة موقع الوسيط .

f₂: التكرار المتجمع اللاحق لقيمة موقع الوسيط.

H: طول (مدى) الفئة.

مثال (7.4): استخدم جدول التوزيع التكراري للمثال (3.4) لإيجاد قيمة الوسيط.

الحل (7.4):

أ. نستخرج التكرار المتجمع الصاعد.

| الفئات | التكرار f_i | التكرار المتجمع الصاعد |
|--------|-----------------|------------------------|
| 40-49 | 3 | 3 |
| 50-59 | 5 | 8 |
| 60-69 | 11 | 19 |
| 70-79 | 6 | 25 |
| 80-89 | 4 | 29 |
| 90-99 | 2 | 31 |
| | $\sum f_i = 31$ | |

ب. نحدد موقع الوسيط:

$$\frac{\sum f_i}{2} = \frac{31}{2} = 15.5$$

وعند النظر الى عمود التكرارات المتجمعة الصاعدة، نجد ان موقع الوسيط

يقع بين القيمتين 8 و 19

ج. نحدد الفئة الوسيطة:

وحيث ان موقع الوسيط هو بين قيمتين، فتكون الفئة المقابلة للتكرار اللاحق

لموقع الوسيط هي الفئة الوسيطة، وبذلك ستكون الفئة (60-69).

د. نطبق الصيغة (11.4) فنحصل على:

$$Md = L + \frac{\frac{\sum f_i}{2} - f_l}{f_2 - f_l} H$$

$$= 60 + \frac{15.5 - 8}{19 - 8} (10)$$

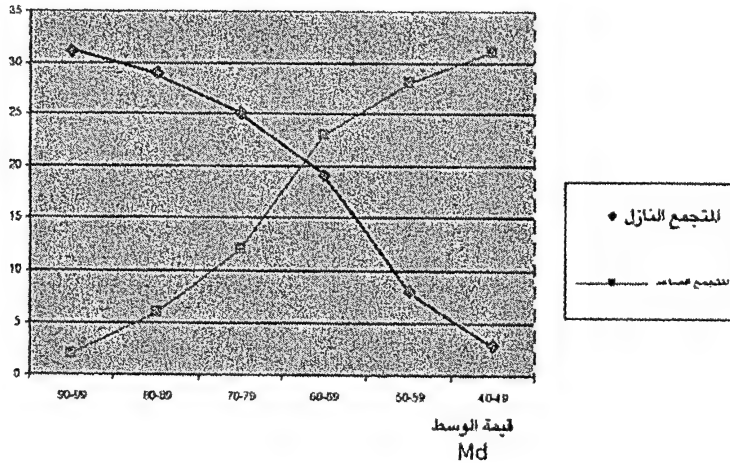
$$60 + 6.8 = 66.8$$

ثانياً: الطريقة البيانية

ويتم ذلك إما من خلال رسم المنحنيين المتجمعين الصاعد والنازل، ومن ثم إنزال خط عمودي من نقطة التقاء المنحنيين على المحور الأفقي، حيث إن نقطة الالتقاء ستمثل موقع الوسيط، والنقطة التي سيقع عليها الخط العمودي على المحور الأفقي ستمثل قيمة الوسيط، كما هو مبين في الشكل البياني (1.4) باستخدام معطيات المثال (4.4). أو الاكتفاء برسم أحد المنحنيين، إما المتجمع الصاعد أو المتجمع النازل وذلك بتحديد موقع الوسيط على المحور العمودي والتوصيل بين الموقع والمنحنى بخط مستقيم، ومن ثم إنزال خط مستقيم من نقطة الالتقاء بالمنحنى إلى المحور الأفقي لتمثل قيمة الوسيط.

شكل بياني رقم (1.4)

استخدام المنحنيين المتجمعين الصاعد والنازل لتحديد قيمة الوسيط



خواص الوسيط وعيوبه

يمتاز الوسيط بالخصائص التالية:

- أولاً: عدم تأثرة بصورة مباشرة بالقيم المتطرفة (أو الشاذة) في تمثيله للمعطيات.
- ثانياً: إمكانية استخدام مع الفئات المفتوحة وغير المتساوية في الطول.
- ثالثاً: سهولة استخراج .

اما عيوبه فتتمثل بـ :

- أولاً: اذا كان عدد المعطيات قليلاً، فالوسيط ممكن أن لا يعبر بصورة صحيحة عن مركز تجمع المعطيات.
- ثانياً: اعتماده على قيمة واحدة او قيمتين في حالة المعطيات غير المبوبة، او على فئة واحدة في حالة المعطيات المبوبة، ولا يأخذ القيم الأخرى بنظر الاعتبار، لذا فإنه يكون حساساً للقيم الوسيطة.
- ثالثاً: قد يتطلب لعمليات غير جبرية مطولة وخاصة في حالة المعطيات غير المبوبة، كترتيب المعطيات تصاعدياً او تنازلياً.

3 المنوال: The Mode

1- حالة المعطيات غير المبوبة، Ungrouped Data:

المنوال هو القيمة الأكثر تكراراً بين مجموعة القيم، ويمكن استخدامه للقيم الكمية والنوعية، وطبقاً لذلك فإن قيمة لا تكون الوحيدة فقد تكون هناك أكثر من قيمة منوالية واحدة، وكل منها أو بعضها يتكرر لعدة مرات، كما هو في حالة مجموعة القيم التالية : 65,65,61,70,78,78,56,56,80,65,56,61,64

فنجد ان كلاً من القيم 78,65,56 قد تكرر وقوعها ثلاث مرات، وعليه فإن هناك ثلاث قيم للمنوال. كما قد لا توجد قيمة منوالية بين القيم، ويحصل ذلك عندما تكون كافة القيم لها نفس العدد من التكرارات. وبصورة عامة يمكن القول إن أكثر استخدامات المنوال يكون مع المعطيات النوعية. فيتم بواسطة التعبير عن صفة

الشيوع، فيقال إن النموذج أو الموديل كذا من الإنتاج هو الأكثر شيوعاً من خلال تكرار مبيعاته أكثر من النماذج الأخرى وهكذا.

مثال (8.4): المعطيات التالية تمثل قيم تبرعات أحد المناطق السكنية (بالدينار) والمطلوب تحديد المنوال. 9,10,5,9,9,7,8,6,10,11

الحل (8.4):

من ملاحظة عدد حالات تكرار كل من القيم أعلاه، نجد أن الرقم (9) قد تكرر ثلاث مرات، في حين تراوحت عدد تكرارات القيم الأخرى بين تكرار واحد وتكرارين. لذا فإن المنوال هو القيمة 9 .

2- حالة المعطيات المبوبة: Grouped data

أولاً: يتم تحديد الفئة المنوالية والتي هي الفئة التي يقابلها أكبر تكرار
ثانياً: نطبق الصيغة التالية:

$$Mo = L + \frac{d_1}{d_1 + d_2} H \quad (13.4)$$

حيث إن :

L: الحد الأدنى للفئة المنوالية.

d_1 : تكرار الفئة المنوالية - تكرار الفئة السابقة.

d_2 : تكرار الفئة المنوالية - تكرار الفئة اللاحقة.

H: طول الفئة.

مثال (9.4): استخدم جدول التوزيع التكراري للمثال (4.4) لإيجاد قيمة المنوال:

الحل (9.4) : لدينا :

| الفئات | التكرار fi |
|---------|------------|
| 40-49 | 3 |
| 50-59 | 5 |
| 60-69 | 11 |
| 70-79 | 6 |
| 80-89 | 4 |
| 90-99 | 2 |
| المجموع | 31 |

وحيث إن أكبر تكرار والبالغ 11 هو للفئة 60-69، لذلك فهي تعتبر الفئة المنوالية، وبتطبيق الصيغة (13.4) نحصل على :

$$Mo = L + \frac{d_1}{d_1 + d_2} H$$

حيث أن :

$$\begin{aligned} 60 &= L \\ 11-5 &= d_1 \\ 11-6 &= d_2 \\ 10 &= H \end{aligned}$$

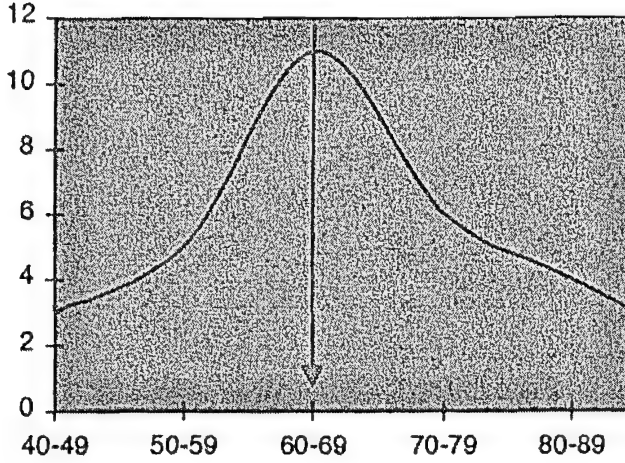
$$Mo = 60 + \frac{6}{5+6} (10)$$

$$= 65.45 \quad \text{قيمة المنوال}$$

ثانياً: الطريقة البيانية

وبواسطتها يمكن إيجاد قيمة المنوال من خلال انزال خط عمودي من قمة المنحنى التكراري على المحور الأفقي، فالنقطة التي يقطعها هذا الخط العمودي تمثل قيمة المنوال. فباستخدام معطيات المثال (9.4) نحصل على الشكل البياني رقم (4.4) ادناه :

شكل بياني رقم (4.4)
إيجاد المنوال بالطريقة البيانية



كما يمكن أيضا إيجاد المنوال باستخدام المدرج التكراري، وذلك بربط زوايا أعلى مضلع تكراري قطريا بزوايا المضلعات المجاورة له، وإنزال خط عمودي من نقطة التقاء الخطوط القطرية على المحور الأفقي لتكون النقطة التي يتقاطع معها على المحور الأفقي هي قيمة المنوال، مع الإشارة بأن استخدام هذه الحالة تنطبق مع المدرج التكراري ذي الفئات المتساوية.

3- خواص المنوال وعيوبه: Mode Properties

أولاً: عدم تأثرة بالقيم المتطرفة (أو الشاذة).

ثانياً: أنه يمثل غالبية المشاهدات.

ثالثاً: احتسابه لا يحتاج لكافة قيم التوزيع.

رابعاً: امكانية احتسابه في حالة الجداول التكرارية ذات الفئات المفتوحة.

وابرز عيوبه تظهر عندما تكون القيم منتشرة على مدى واسع، عندها يصبح أقل تعبيراً كمتوسط.

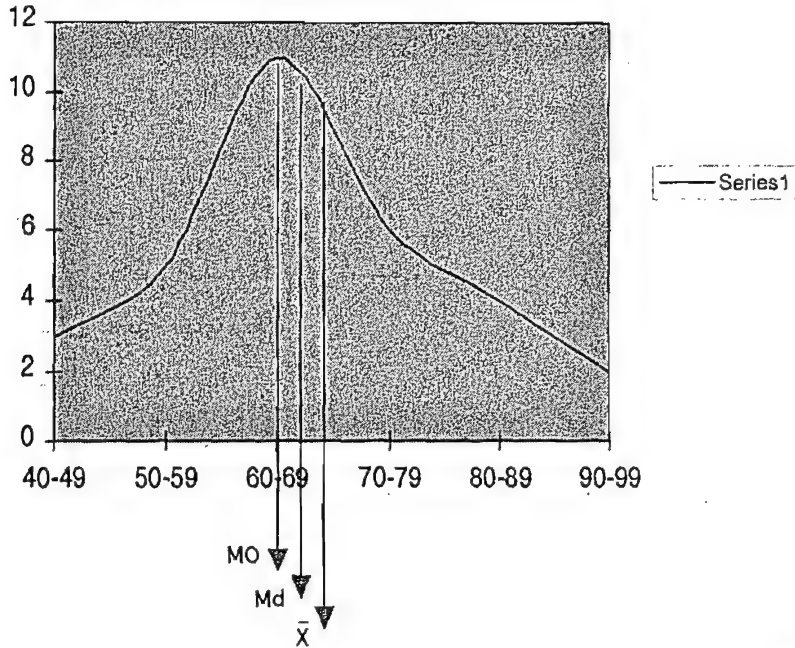
4. العلاقة التقريبية بين الوسط الحسابي والوسيط والمنوال

Approximate Relation of the Mean, Median, and Mode :

مما سبق يكمن الاستنتاج بأن الوسط الحسابي يقسم بصورة متساوية المساحة تحت المنحنى الى مجموع الانحرافات السالبة على الجانب الأيسر ومجموع الانحرافات الموجبة على الجانب الأيمن. فهو بذلك يمر من النقطة المركزية للمساحة تحت المنحنى. وان الوسيط يقسم المساحة تحت المنحنى الى قسمين متساويين بحيث أن عدد المعطيات التي تقل عن قيمة الوسيط مساوية لعدد المعطيات التي تقل عن قيمة الوسيط. بينما قيمة المنوال تطابق أعلى نقطة على المنحنى. ويمكن تصور هذه العلاقة لمقاييس النزعة المركزية الثلاثة بالشكل البياني رقم (4.5) التالي :

شكل بياني رقم (5.4)

يوضح العلاقة بين قيم الوسط الحسابي والوسيط والمنوال عندما يكون التفرطح باتجاه اليمين



ويتطابق الوسط الحسابي والوسيط والمنوال فقط عندما يكون شكل المنحنى متماثلاً ومتجانساً تماماً (Symmetric) أما في حالة عدم تحقق هذا التماثل، فإن المنحنى يقال عنه مفرطح (Skewness)، فعندما يكون التفرطح باتجاه اليمين كما في الشكل (4-5) السابق، سيكون المنوال إلى يسار الوسيط، والوسط الحسابي على يمينه، أما عندما يكون التفرطح باتجاه اليسار فسيكون المنوال إلى يمين الوسيط والوسط الحسابي.

أما في حالة التوزيعات التي يكون الالتواء فيها معتدلاً، فإن العلاقة التقريبية بين المتوسطات الثلاثة تصبح كالآتي :

الوسط الحسابي - المنوال = 3 (الوسط الحسابي - الوسيط)

5. الوسط الهندسي : The Geometric Mean

ويستخدم هذا النوع من المتوسطات مع النسب ومعدلات النمو ومع الأرقام القياسية.

1- حالة المعطيات غير المبوبة، Ungrouped Data :

يعرف الوسط الهندسي بأنه جذر n لقيم عددها n فإذا رمزنا له بـ \bar{X}_g فإن الوسط الهندسي لعينة حجمها n وقيمها هي X_1, X_2, \dots, X_n سيكون عبارة عن الصيغة

$$\bar{X}_g = \sqrt[n]{(X_1)(X_2) \dots (X_n)} \quad (15.4)$$

وبتحويل قيم المتغير X إلى $\log x$ فإن الصيغة (15.4) تصبح :

$$\log \bar{X}_g = \frac{1}{n} \sum \log X_i$$

وان قيمة اللوغارتم المقابل (antilogarithm) للنتيجة تمثل الوسط الهندسي.

مثال: (10.4): أوجد الوسط الهندسي للقيم التالية: 1.2، 1.5، 1.67، 2.0، 1.67

الحل (10.4):

باستخدام الصيغة (15.4) يكون لدينا :

$$\bar{X}_g = \sqrt[5]{(1.67)(2.0)(1.67)(1.5)(1.2)}$$

$$\begin{aligned} \log \bar{X}_g &= \frac{1}{5} \sum \log (1.67)(2.0) \dots \dots \dots (1.2) \\ &= \frac{1}{5} (0.2227 + 0.301 + 0.2227 + 0.1761 + 0.097) \\ &= \frac{1}{5} (1.0101) = 0.20202 \end{aligned}$$

وبإيجاد اللوغاريتم المقابل (10^x) نحصل على :

$$\bar{X}_g = 1.5923$$

في حين عند استخدام الوسط الحسابي \bar{X} مع قيم المثال (10.4) نحصل على $\bar{X} = 1.61$ ، ان سبب الاختلاف يعود الى تأثير الوسط الحسابي بتباين حجم القيم. وللزيادة في التوضيح لو تأملنا بالقيمة 100 تهبط الى 50 ومن ثم ترتفع الى 100 فإن مقدار التغير هو 0.5, 2.0 على التوالي وبذلك فإن الوسط الهندسي سيكون :

$$\bar{X}_g = \sqrt{(0.5)(2.0)}$$

$$\log \bar{X} = \frac{1}{2} \sum \log(0.50)(2.0)$$

وبإيجاد اللوغاريتم المقابل نحصل على: $\bar{X}_g = 1$

$$\bar{X} = \frac{0.5 + 2.0}{2} = 1.25 \quad \text{لكن عند استخدام الوسط الحسابي سيكون لدينا}$$

وهي نتيجة غير واقعية طبقا للسبب أعلاه.

مثال (11.4): أوجد الوسط الهندسي للارقام القياسية التالية لاسعار الجملة ل 8 مجموعات سلعية لشهر تشرين الثاني 1992.

101، 90، 108، 96، 103، 79، 85، 100

الحل (11.4):

$$\log 101 = 2.0043 \text{ لدينا}$$

$$\log 90 = 1.9956$$

$$\log 108 = 2.0334$$

$$\log 96 = 1.9823$$

$$\log 103 = 2.0128$$

$$\log 79 = 1.8976$$

$$\log 85 = 1.9294$$

$$\log 100 = 2.000$$

$$\sum \log X_i = 15.8554$$

وبالتعويض بالصيغة (15.4) نحصل على:

$$\log \bar{X} = \frac{15.8554}{8} = 1.9819$$

وباستخراج القيمة (10^x) لـ \bar{X}_g نجد أن الوسط الهندسي هو: $\bar{X}_g = 96$

وهنا يجدر التنكير من أن استخدام الوسط الهندسي سيكون فقط مع القيم الموجبة، حيث لا يمكن استخدام مع القيم السالبة أو الصفر.

2- في حالة المعطيات المبوبة Grouped data

$$\bar{X}_g = \sqrt[n]{(X_1^{f_1})(X_2^{f_2}) \dots (X_n^{f_n})} \quad (16.4)$$

وباستخدام اللوغاريتم تصبح الصيغة (16.4) :

$$\log \bar{X}_g = \frac{1}{\sum f_i} \sum f_i \log X_i$$

مثال (12.4) : أوجد الوسط الهندسي لقيم جدول التوزيع التكراري للمثال (4.4).

الحل (12.4) :

نجد قيم كل من $\log x_i$ ، $f_i \log x_i$ وكالاتي:

| الفئات | التكرار (f_i) | X_i | $\log X_i$ | $f_i \log X_i$ |
|--------|-------------------|-------|------------|----------------------------|
| 40-49 | 3 | 44.5 | 1.648 | 4.944 |
| 50-59 | 5 | 54.5 | 1.736 | 9.68 |
| 60-69 | 11 | 64.5 | 1.809 | 19.899 |
| 70-79 | 6 | 74.5 | 1.872 | 11.232 |
| 80-89 | 4 | 84.5 | 1.926 | 7.704 |
| 90-99 | 2 | 94.5 | 1.975 | 3.95 |
| | $\sum f_i = 31$ | | | $\sum f_i \log X = 57.409$ |

وبتطبيق الصيغة (16.4) نحصل على:

$$\log \bar{X}_g = \frac{1}{\sum f_i} \sum f_i \log X_i$$

$$= \frac{57.409}{31} = 1.852$$

وبإعادة اللوغاريتم المقابل (10^x) نحصل على الوسط الهندسي وهو:

$$\bar{X}_g = 71.12$$

6. الوسط التوافقي Harmonic mean

ويتركز استخدام الوسط التوافقي (\bar{X}_h) في الغالب عندما يراد إيجاد المتوسط وفقا لوحدة قياسية معينة كالذينة او الصندوق الذي يحتوي على عدد معين من القناني او العلب وما شابه.

1- حالة البيانات غير المبوبة ungrouped data

ويكون عبارة عن مقلوب (reciprocals) الوسط الحسابي، أي:

$$\bar{X}_h = \frac{1}{\sum \frac{1}{X_i}}$$

مثال (4-13): إذا كانت نفقات الأسرة على شراء البيض هو 3 دينار شهريا وكان سعر الدرزن الواحد من البيض هو 0.800 دينار، وفي الشهر اللاحق انفقت ذات الأسرة 4 دنانير على البيض وكان سعر الدزينة الواحدة 1.100 دينار، فما هو متوسط سعر الدزينة الواحدة.

الحل (4-13): باستخدام الوسط التوافقي اعلاه يكون لدينا :

$$\bar{X}_h = \frac{2}{\frac{1}{0.800} + \frac{1}{1.100}} = 0.926$$

2- حالة البيانات المبوبة grouped data

$$\bar{X}_h = \frac{\sum f_i}{\sum \left\{ \frac{f_i}{X_i} \right\}}$$

وفي هذه الحالة تصبح الصيغة:

مثال (4-14): اوجد الوسط التوافقي للمثال (4-4).

| الفئات | التكرار (f_i) | X_i | $\frac{f_i}{x_i}$ |
|--------|-------------------|-------|-------------------|
| 40-49 | 3 | 44.5 | 0.067 |
| 50-59 | 5 | 54.5 | 0.092 |
| 60-69 | 11 | 64.5 | 0.171 |
| 70-79 | 6 | 74.5 | 0.081 |
| 80-89 | 4 | 84.5 | 0.047 |
| 90-99 | 2 | 94.5 | 0.021 |
| | $\sum f_i = 31$ | | 0.479 |

وبتطبيق صيغة الوسط التوافقي اعلاه نحصل على الوسط التوافقي المبين في

ادناه:

$$\bar{X}_h = \frac{31}{0.479} = 64.718$$

3- خواص الوسط التوافقي وعيوبه

من ابرز عيوب هذا المقياس تأثره بالقيم المتطرفة في الصغر، ويصبح ليس ذا مدلول مع وجود هذا النوع من القيم، ويفضل استخدامه عند البحث عن متوسط التغير عبر الزمن.

4.4 مقاييس التشتت (التباين) Measures Of Variation

وتتناول كيفية قياس انتشار البيانات حول نقطة التركيز (المتوسط)، فمن الممكن جدا ان يكون لمجموعتين من البيانات نفس المتوسط وان يكونا مختلفين معنويا في انتشارهما حول المتوسط، فلو تأملنا في المثال التالي الذي يمثل عدد افراد عينتين من الأسر وهي:

| عدد أفراد الأسرة للعينة الثانية | عدد أفراد الأسرة للعينة الأولى |
|------------------------------------|-----------------------------------|
| 2 | 4 |
| 6 | 5 |
| 1 | 6 |
| 11 | 5 |

نجد ان الوسط الحسابي لكلا العينتين متساويين: $\bar{x}_1 = \bar{x}_2 = 5$ رغم التباين الواضح في عدد افراد كلا الأسرتين.

1. المدى Range

1- حالة المعطيات غير المبوبة Ungrouped Data:

والمدى هو عبارة عن الفرق بين اكبر قيمة واصغر قيمة بين البيانات المعنية، فالمدى لبيانات العينة الاولى في المثال اعلاه هو: $R = 6 - 4 = 2$ بينما المدى للعينة الثانية هو: $R = 11 - 1 = 10$. وبرغم سهولة وبساطة حساب المدى الا انه يعتبر من مقاييس التشتت غير الدقيقة، لانه يعتمد على القيم المتطرفة فقط وإهمال بقية القيم بينهما. ان اغلب استخداماته هو في مجال السيطرة النوعية للإنتاج وفي مجال قياس التغير في درجات الحرارة .

2 - حالة المعطيات المبوبة Grouped Data

بالنظر لمجهولية اصغر واكبر قيمة في حالة البيانات المبوبة، فإن قيمة المدى التقديرية تكون عبارة عن الفرق بين الحد الاعلى للفئة العليا والحد الادنى للفئة الدنيا، فالمدى للمثال رقم (4-4) هو: $R = 99 - 40 = 59$.

2 الانحراف المعياري Standard Deviation

1- حالة المعطيات غير المبوبة Ungrouped Data:

ويعتبر الاكثر اهمية واستخداما كمقياس للتشتت، ويرمز له في حالة العينة ب s وفي حالة المجتمع ب σ . وان صيغة احتسابه في حالة البيانات غير المبوبة هي :

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}}$$

ويمكن تلخيص العمليات الحسابية المطلوبة لإيجاد الانحراف المعياري بالخطوات التالية :

1. استخراج الوسط الحسابي للبيانات.
2. ايجاد انحرافات القيم x_i عن الوسط الحسابي \bar{X} .
3. تربيع كل انحراف من الانحرافات المعنية.
4. ايجاد مجموع مربعات الانحرافات.
5. تقسيم مجموع مربعات الانحرافات على عدد القيم $n-1$ فنحصل على التباين S^2 .
6. اخذ الجذر التربيعي للتباين نحصل على الانحراف المعياري s .

مثال (4-15): اوجد الانحراف المعياري s لقيم المشاهدات التالية: 7, 3, 12, 8, 5

الحل (4-15): باتباع الخطوات اعلاه، لدينا :

$$\bar{X} = \frac{\sum X_i}{n} = \frac{35}{5} = 7$$

$$\sum (x_i - \bar{x}) = -2 + 1 + 5 + -4 + 0$$

$$\sum (x_i - \bar{x})^2 = 4 + 1 + 25 + 16 + 0 = 46$$

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

$$= \sqrt{11.5} = 3.391$$

وبالامكان اختصار العمليات الحسابية اعلاه، باستخدام الصيغة التالية وهي عبارة عن مفكوك للصيغة السابقة (في حالة الرغبة في التفصيل يمكن الرجوع الى كتاب "الاحصاء للعلوم الادارية والتطبيقية" للمؤلف، 1997).

$$s = \sqrt{\frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n - 1}}$$

وبتطبيق الصيغة اعلاه على المثال (4-15) نحصل على:

$$\sum x_i = 35$$

$$\sum x_i^2 = 49 + 9 + 144 + 64 + 25 = 291$$

$$s = \sqrt{\frac{\sum 291 - \frac{(35)^2}{5}}{4}} = 3.391$$

2- حالة البيانات المبوبة grouped data

وفي هذه الحالة نقوم اولا بايجاد مراكز الفئات x_i ومن ثم حساب مربعاتها x_i^2 ، وباستخدام الصيغة التالية نحصل على الانحراف المعياري s :

$$s = \sqrt{\frac{\sum f_i x_i^2 - \frac{(\sum f_i x_i)^2}{n}}{n - 1}}$$

حيث ان: $n = \sum f_i$

مثال (4-16): لدينا جدول التوزيع التكراري موضوع المثال (4-4)، والمطلوب إيجاد الانحراف المعياري s .

الحل (4-16):

| الفئات | f_i | x_i | $f_i x_i$ | X_i^2 | $\sum f_i X_i^2$ |
|--------|-----------------|-------|-------------------------|---------|------------------------------|
| 40- 49 | 3 | 44.5 | 133.5 | 1980.25 | 5940.75 |
| 50- 59 | 5 | 54.5 | 272.5 | 2970.25 | 14851.25 |
| 60- 69 | 11 | 64.5 | 709.5 | 4160.25 | 45762.75 |
| 70- 79 | 6 | 74.5 | 447 | 5550.25 | 33301.5 |
| 80- 89 | 4 | 84.5 | 338 | 7140.25 | 28561 |
| 90- 99 | 2 | 94.5 | 189 | 8930.25 | 17860.5 |
| | $\sum f_i = 31$ | | $\sum f_i x_i = 2089.5$ | | $\sum f_i x_i^2 = 146277.75$ |

وبتطبيق الصيغة اعلاه نحصل على:

$$s = \sqrt{\frac{\sum f_i x_i^2 - \frac{(\sum f_i x_i)^2}{n}}{n - 1}}$$

$$s = \sqrt{\frac{\sum 146277.75 - \frac{(2089.5)^2}{31}}{30}} = 4.9$$

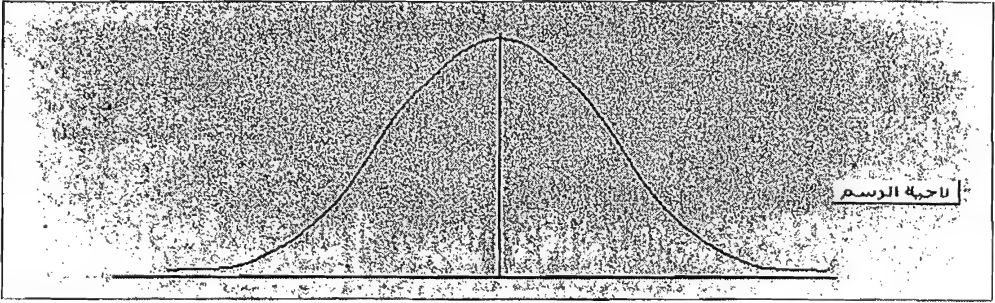
ورغم انتهاء الحاجة بعد شيوع استخدام الحاسوب الى الاختصار في العمليات الحسابية، فهناك اكثر من طريقة للاختصار (في حالة الرغبة في التفصيل يمكن الرجوع لكتاب "الاحصاء للعلوم الادارية والتطبيقية" للمؤلف).

5.4 مقاييس التماثل والالتواء.

Symmetric and Skewness Measures □

يقال بان التوزيع متماثل عندما يتطابق نصفاً شكل التوزيع الطبيعي عند محور عمودي كما هو مبين في الشكل البياني (4-6) التالي، لكن عندما لا يتطابق جانبا التوزيع يقال عنه ملتوياً، فعندما يكون الالتواء باتجاه اليمين يقال عنه توزيع موجب الالتواء، ويحصل ذلك اذا كان الوسط الحسابي يزيد على الوسيط، اما عندما يكون الالتواء باتجاه اليسار فعندها يدعى بالتوزيع سالب الالتواء، وهو الحالة التي يقل فيها الوسط الحسابي عن الوسيط

الشكل رقم (4-6)
شكل التوزيع المتماثل



ومن اهم مقاييس التماثل والالتواء ولنرمز له Sk هو معامل بيرسن Pearsnian Coefficient وصيغته:

$$Sk = \frac{3(\bar{x} - Md)}{S}$$

وبصورة عامة فان قيمة الالتواء تقع بين $Sk = \pm 3$ وتصبح قيمته 0 في حالة تطابق قيم المتوسطات.

مثال (4-17): احسب معامل بيرسن للالتواء لتوزيع علامات الطلبة في مادة الاحصاء
موضوع المثال (3-1).

الحل (4-17):

$$\text{لدينا: } \bar{x} = 67.4, \quad Md = 66.8, \quad s = 4.9$$

وبتطبيق صيغة معامل بيرسن Sk نحصل على :

$$Sk = \frac{3(\bar{x} - Md)}{S}$$

$$Sk = \frac{(66.8) - 3(67.4)}{4.9} = 0.367$$

ومن النتيجة نستدل على ان الالتواء موجب بسيط يمكن التعبير عنه بأنه قريب
للتماثل، كما هو مبين من الشكل البياني رقم (4-4).

تمارين الفصل الرابع

تمرين (4-1): في اختبار المعلومات على 28 من طلبة ادارة الاعمال وعند 10

درجات، كانت حصيلة الاختبار هي كما مبين في التالي، والمطلوب:

أ. استخدام برنامج SPSS لايجاد كل من: الوسط الحسابي، الوسيط،

المنوال، معامل الالتواء، الانحراف المعياري

ب. ايجاد المقاييس في الفقرة أ بالطريقة اليدوية.

8، 6، 10، 3، 4، 7، 8، 6، 4، 8، 6، 4، 5، 5، 3، 1، 7، 8، 5،

2، 4، 5، 6، 9، 5، 6

تمرين (4-2): استخدم الطريقة اليدوية مع جدول التوزيع التكراري التالي الذي

يمثل عدد القروض المقدمة من قبل احد البنوك موزعة حسب فئات

مبالغ القروض (بالدينار) لايجاد :

أ. مقاييس النزعة المركزية بالطريقة الحسابية.

ب. مقاييس النزعة المركزية بالطريقة البيانية.

ج. مقاييس التشتت.

| الفئات | عدد القروض f_i |
|-----------|------------------|
| 199 فأقل | 6 |
| 200-399 | 18 |
| 400-599 | 25 |
| 600-799 | 20 |
| 800-999 | 17 |
| 1000-1199 | 14 |
| المجموع | 100 |

تمرين (3-4): البيانات التالية نسبة الوفيات بسبب حوادث الطرق لكل 100 مليون (كم/واسطة نقل) لعدد من محافظات احدى الدول لسنة 2002، والمطلوب إيجاد الوسط الهندسي:

| النسبة | المحافظة |
|--------|----------|
| 4.1 | A |
| 1.3 | B |
| 3.7 | C |
| 4.4 | D |
| 4.7 | E |
| 1.3 | F |
| 4.0 | G |
| 2.4 | H |

تمرين (4-4): استخدم مقياس بيرسن مع تمرين (2-4) لقياس التماثل والالتواء .

تمرين (5-4): وضح خصائص الوسط الحسابي، وبماذا يتميز عن باقي مقاييس النزعة المركزية .



الارتباط
CORRELATION

1-5 - مقدمة

يستهدف الارتباط معرفة ان كانت هناك علاقة بين متغيرين او مجموعة متغيرات مستقلة X_i والمتغير التابع Y ، وهناك مقياسان لتحديد درجة الارتباط هما: معامل الارتباط ونرمز له r في البحث عن العلاقة بين متغيرين و نرمز له R عند البحث عن العلاقة بين المتغير التابع مع متغيرين مستقلين فاكثراً. والمقياس الثاني هو معامل التحديد Coefficient of Determination والذي هو عبارة عن مربع معامل الارتباط. ويقال ان الارتباط موجب اذا كانت قيم المتغير التابع Y تميل الى الارتفاع كلما ارتفعت قيم X ، اما اذا كانت قيم Y تميل نحو الانخفاض كلما ارتفعت قيم X فيقال بان الارتباط سالب. ومن خواص معامل الارتباط ان قيمته تقع بين 0 و 1، فعندما $r=0$ فتعني عدم وجود اي نوع من الارتباط .

2.5 استخدام الحاسوب مع برنامج SPSS

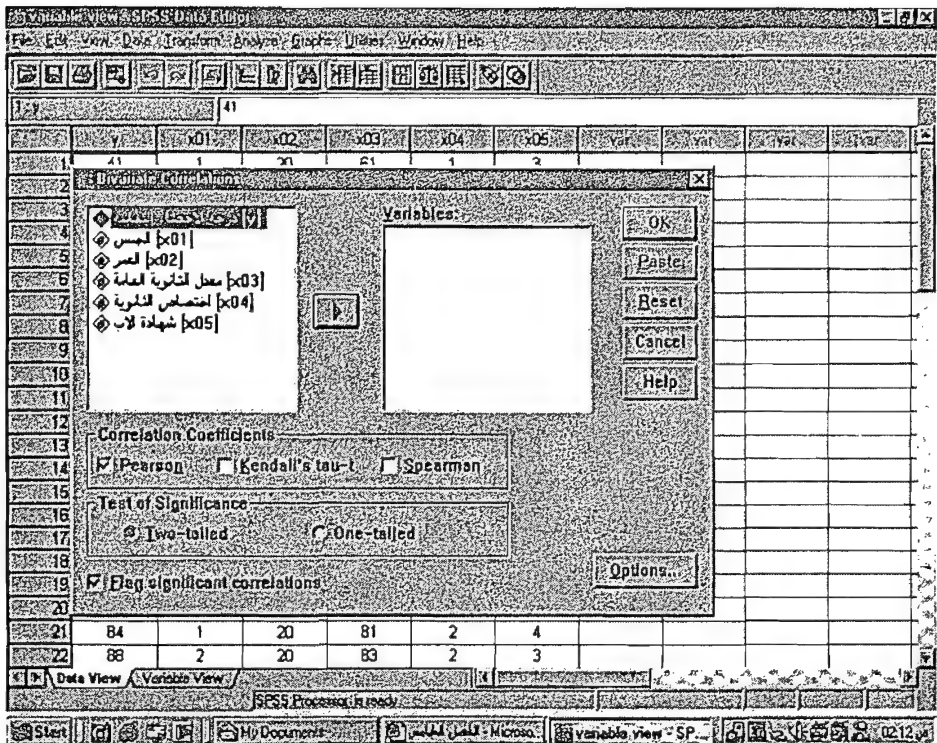
هناك عدة انواع من الارتباط، استخدام كل منها يتم إما حسب الحاجة الى استبعاد تأثير متغيرات معينة او الابقاء عليه، وايضا حسب طبيعة البيانات (كمية او نوعية)، فاذا كانت البيانات كمية يستخدم لقياسه معامل ارتباط Pearson. اما اذا كانت البيانات نوعية (غير رقمية) فيمكن استخدام معامل ارتباط Spearman او Kendall's tau-t وهذه الانواع الثلاثة تدخل ضمن الامر Bivariate من الامر الفرعي Correlate من خلال الامر الرئيسي Analyze، اما النوعان الآخران المتوفران ضمن الامر الفرعي Correlate فهما Partial ويستخدم عند الحاجة لاستبعاد تأثير المتغيرات الاخرى على علاقة متغيرين محددين، والنوع الآخر هو Distance واستخدامه لغرض الوقوف على العلاقة بين المشاهدات cases او الاقتصار على العلاقة بين المتغيرات Variables. وبعد الحصول على المخرجات Output والتي تكون عادة على شكل مصفوفة ستظهر على معاملات الارتباط اشارة تحمل شكل نجمة * او نجمتين ** لتدل الاولى على درجة المعنوية (الدلالة) عند 0.05 اي معنوية، في حين تدل النجمتان على درجة المعنوية عند 0.01 اي عالية المعنوية.

وباستخدام بيانات المثال (1.3) مع معامل ارتباط Pearson مثلاً نحتاج الى الاجراءات التالية، وعلى افتراض قد تم دخولنا الى برنامج SPSS وتاشيرنا على ملف البيانات المتعلق بعينة الطلبة موضوع مثالنا المستهدف اخضاعه لعملية التحليل:

Analyze → Correlate → Bivariate •

• وبعد التاشير على الامر Pearson والكبس (click) سيظهر لنا مربع الحوار ومتضمن المتغيرات الموجودة في الملف كما هو مبين في الشكل رقم (1-5) :
 1-5

شكل رقم (1.5)
 مربع حوار الامر Bivariate

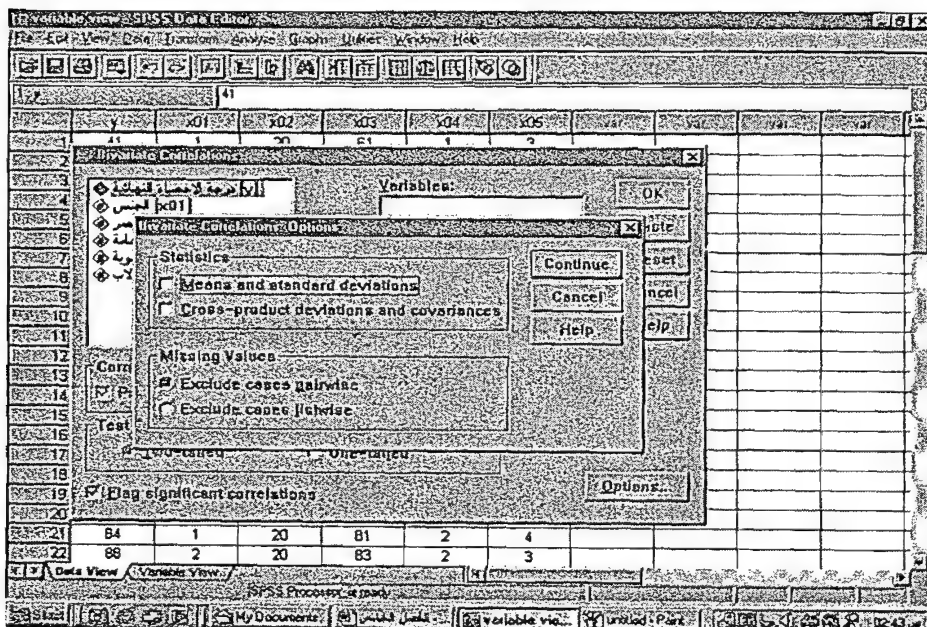


• القيام بتحديد المتغيرات المطلوب ايجاد العلاقات بينها من خلال الكبس على السهم المبين على يمين المربع المتوفرة فيه المتغيرات لنتنقل الى المربع الآخر الى الجانب الايمن .

- اختيار نوع الارتباط المطلوب استخدامه في عملية التحليل والمبينة انواعها في اسفل المربعات فنؤشر على Pearson بالنسبة لمثالنا .
- التأشير على احد الخيارين المتعلقة باختبار من جانب واحد One tail اذا كانت هناك معرفة مسبقة باتجاه العلاقة او جانبيين Two tail في حالة عدم المعرفة المسبقة باتجاه العلاقة، ويفضل في الغالب التأشير على جانبيين Two tail ضممانا للدقة وتلافيا للخطورة.
- الكبس على ايقونة Option المبينة في الزاوية اليمنى عند اسفل مربع الحوار للحصول على مربع حوار اضافي، والمبين نموذجه في الشكل البياني رقم (2-5) في حالة الرغبة للحصول على الوسط الحسابي او الانحراف المعياري لكل من المتغيرات تحت التحليل. ويتم الرجوع من مربع الحوار الاضافي الى مربع الحوار الاساسي بالكبس على ايقونة Continue.

شكل رقم (25)

يوضع مربع الحوار الاضافي باستخدام الايقونة Option من مربع حوار الامر Bivariate



• وبعد العودة الى مربع الحوار الاساسي يتم الكبس على ايقونة OK للحصول على المخرجات المبين نموذجها في الجدول رقم (5-1) .

جدول رقم (15)

بوضع مخرجات Output تحليل الارتباط بطريقة Pearson

Correlations

| شهادة الأب | اختصاص ثانوية | معدل الثانوية العامة | العمر | الجنس | درجة الإحصاء قنهائية | |
|---------------|------------------|-------------------------|-------|---------|-------------------------|---------------------|
| .242 | .348 | .601** | .233 | -.080 | 1.000 | Pearson Correlation |
| .189 | .055 | .000 | .208 | .688 | 0 | Sig.(2-tailed) |
| 31 | 31 | 31 | 31 | 31 | 31 | N |
| -.469** | -.172 | -.014 | -.040 | 1.000 | -.080 | Pearson Correlation |
| .008 | .354 | .939 | .829 | 0 | .668 | Sig.(2-tailed) |
| 31 | 31 | 31 | 31 | 31 | 31 | N |
| .082 | .035 | .115 | 1.000 | -.040 | .233 | Pearson Correlation |
| .662 | .851 | .538 | 0 | .829 | .208 | Sig.(2-tailed) |
| 31 | 31 | 31 | 31 | 31 | 31 | N |
| .460** | .578** | 1.000 | .115 | -.014 | .601** | Pearson Correlation |
| .009 | .001 | 0 | .538 | .939 | .000 | Sig.(2-tailed) |
| 31 | 31 | 31 | 31 | 31 | 31 | N |
| .317 | 1.000 | .578** | .035 | -.172 | .348 | Pearson Correlation |
| .082 | 0 | .001 | .851 | .354 | .055 | Sig.(2-tailed) |
| 31 | 31 | 31 | 31 | 31 | 31 | N |
| 1.000 | .317 | .460** | .082 | -.469** | .242 | Pearson Correlation |
| 0 | .082 | .009 | .662 | .008 | .189 | Sig.(2-tailed) |
| 31 | 31 | 31 | 31 | 31 | 31 | N |

** Correlation is significant at the 0.01 level (2-tailed).

ومن بين ما يمكن الاستدلال عليه من المصفوفة اعلاه هو ان المتغير التابع y (علامات مادة الاحصاء) على علاقة قوية (عند مستوى معنوية 0.01) مع كل من المتغيرات المستقلة التالية:

- معدل الطالب في الثانوية العامة
- مستوى الشهادة الدراسية للاب
- فرع دراسة الطالب في الثانوية العامة (علمي - ادبي)

3.5 الطريقة اليدوية

في حالة البيانات النوعية سيتطلب الامر هنا تبويبها بما يتلاءم واجراء عملية حساب معامل الارتباط المستهدف وكما سنرى ذلك لاحقاً،

1. معامل الارتباط البسيط r

والذي يدعى معامل ارتباط بيرسن Pearson's Correlation Coefficient وهو يخص العلاقة بين متغيرين وليس مهما ايهما يكون المتغير التابع وايهما المستقل. ويمكن التعبير عن صيغة العلاقة كالآتي:

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{\{n \sum x^2 - (\sum x)^2\} \{n \sum y^2 - (\sum y)^2\}}}$$

حيث إن :

x = المتغير المستقل

y = المتغير التابع

n = عدد المشاهدات

مثال (5-1): الجدول التالي يضم معدل الدخل x ومصرفات y لعينة تتكون من 6 أسر. والمطلوب حساب معامل الارتباط r .

| التسلسل | معدل الدخل السنوي x_i (بالآلاف الدنانير) | معدل المصروف السنوي y_i (بالآلاف الدنانير) |
|---------|---|---|
| 1 | 2.4 | 1.9 |
| 2 | 1.8 | 1.7 |
| 3 | 3.1 | 2.7 |
| 4 | 2.8 | 2.6 |
| 5 | 4.6 | 3.9 |
| 6 | 3.2 | 3.1 |
| | $\sum x = 17.9$ | $\sum y = 15.9$ |

الحل (1-5):

نجد قيم كل من $\sum y$ ، $\sum x$ ، $\sum y^2$ ، $\sum x^2$ ، $\sum xy$

| y^2 | x^2 | xy |
|--------------------|--------------------|-------------------|
| 3.61 | 5.76 | 4.56 |
| 2.89 | 3.24 | 3.06 |
| 7.29 | 9.61 | 8.37 |
| 6.76 | 7.84 | 7.28 |
| 15.21 | 21.16 | 17.94 |
| 9.61 | 10.24 | 9.92 |
| $\sum y^2 = 45.37$ | $\sum x^2 = 57.85$ | $\sum xy = 51.13$ |

وبتطبيق صيغة الارتباط البسيط اعلاه نحصل على :

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{\{n \sum x^2 - (\sum x)^2\} \{n \sum y^2 - (\sum y)^2\}}}$$

$$= \frac{6(51.13) - (17.9)(15.9)}{\sqrt{\{6(57.85) - (17.9)^2\} \{6(45.37) - (15.9)^2\}}} = 0.974$$

ولاجل اختبار حجم معامل الارتباط r ، فبالامكان استخدام الصيغة التالية

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

ومقارنة نتيجة الصيغة مع قيمة t الجدولية مع درجات حرية $n-2$ و عند مستوى معنوية α ، فاذا كانت القيمة المحسوبة بموجب الصيغة اعلاه هي اكبر من قيمة t الجدولية، عندها نستدل على معنوية العلاقة بين المتغيرين. فعند التعويض بالصيغة اعلاه باستخدام المثال اعلاه يصبح لدينا :

$$t = 0.974 \sqrt{\frac{4}{0.051}} = 8.625$$

وبالمقارنة مع القيمة الجدولية $t_{0.05, 4} = 2.132$ نستدل على العلاقة المعنوية بين المتغيرين المعنيين .

2- معامل الارتباط المتعدد , R Multiple Correlation Coefficient

وهو يبحث في العلاقة بين أكثر من متغيرين، وهو امتداد لمعامل الارتباط البسيط، إلا أن الأمر يصبح أكثر صعوبة إذا أصبح الأمر يتعلق بأكثر من ثلاثة متغيرات، مما يستوجب اللجوء إلى استخدام الحاسوب. مع التنويه هنا إلى الإشارة السالبة والموجبة هنا لاتدل على الاتجاه لأن الأمر يتعلق بأكثر من متغيرين. أما صيغة احتساب معامل الارتباط المتعدد في حالة لدينا ثلاثة متغيرات هي x_1, x_2, y فهي:

$$R_{y.12} = \sqrt{\frac{r^2y1 + r^2y2 - 2ry1 \quad ry2 \quad r12}{1 - r^212}}$$

وكما يتضح من الصيغة اعلاه، فهي بحاجة إلى إيجاد معاملات الارتباط البسيط لكل من: $ry1 \quad ry2 \quad r12$

مثال (5-2): أرادت إحدى الشركات معرفة العلاقة بين عدد المستجيبين Y لإعلاناتها وبين حجم الاعلان $X1$ (بالسم) في الصحيفة وعدد النسخ الموزعة منها $X2$ ، واستطاعت الشركة الحصول على البيانات التالية :

| $X2$ (عدد النسخ الموزعة من الصحيفة بالاف) | $X1$ (حجم الاعلان بالسم) | Y (عدد المستجيبين بالمئات) |
|---|-----------------------------|------------------------------------|
| 2 | 1 | 1 |
| 8 | 8 | 4 |
| 1 | 3 | 1 |
| 7 | 5 | 3 |
| 4 | 6 | 2 |
| 6 | 10 | 4 |

الحل (5-2): لدينا :

$$\begin{array}{lll} \sum y = 15 & \sum x_1 = 33 & \sum x_2 = 28 \\ \sum y^2 = 47 & \sum x_1^2 = 235 & \sum x_2^2 = 170 \\ \sum yx_1 = 103 & \sum yx_2 = 88 & \sum x_1x_2 = 188 \end{array}$$

وبتطبيق صيغة الارتباط البسيط نحصل على:

$$r_{y1}=0.936 \quad r_{y2}=0.931 \quad r_{12}=0.763$$

وبتطبيق صيغة معامل الارتباط المتعدد يكون لدينا :

$$R_{y.12} = \sqrt{\frac{r^2y1 + r^2y2 - 2ry1 \quad ry2 \quad r12}{1 - r^212}}$$

$$R_{y.12} = \sqrt{\frac{0.876 + 0.866 - 2(0.936)(0.931)(0.763)}{1 - 0.582}} = 0.99$$

اما صيغة اختبار معنوية معامل الارتباط المتعدد فهي :

$$\begin{aligned} F &= \frac{R^2_{y.12} - k}{1 - R^2_{y.12}} \cdot \frac{n - k - 1}{k} \\ &= \frac{0.95}{1 - 0.98} \cdot \frac{2}{3} = 32.667 \end{aligned}$$

وحيث ان قيمة F المحتسبة بموجب صيغة الاختبار اعلاه اكبر من القيمة الجدولية

عند $F_{\alpha=0.025, 3, 2} = 16.04$ نستدل على معنوية المعامل R .

3- معامل الارتباط الجزئي Partial Correlation Coefficient

وهو مقياس لارتباط زوج من المتغيرات عندما باقى المتغيرات تبقى ثابتة، فمثلا اذا كانت معادلة ما تضم المتغيرات x_1, x_2, x_3, x_4 فايجاد الارتباط الجزئي بين المتغيرين x_1, x_2 يتم بابقاء المتغيرين الآخرين في المعادلة ثابتة وهذه هي نقطة

الفرق مع معامل الارتباط البسيط، عندها يرمز لمعامل الارتباط الجزئي $r_{12.34}$ ،
ويستخدم هذا النوع من الارتباط في تحليل الانحدار للحالات التالية :

- لمعرفة طبيعة العلاقة بين متغيرين محددين.
- للوقوف فيما اذا كانت هناك متغيرات يجب حذفها من معادلة الانحدار بسبب محدودية او انعدام تأثيرها على المتغير التابع.
- لاضافة متغير او اكثر الى المعادلة لأجل تحسين قوة وكفاءة المعادلة التنبؤية.

ان صيغة حساب معامل الارتباط الجزئي بين y و x_2 مع ثبات x_1 تأخذ الشكل التالي :

$$r_{y2.1} = \frac{ry2 - (ry1)(r12)}{\sqrt{(1 - r^2y1)(1 - r^212)}}$$

مثال (3-5): المطلوب ايجاد معامل الارتباط الجزئي $r_{y2.1}$ للمثال السابق (2-5) .

الحل (3-5): لدينا:

$$r_{y1} = 0.936 \quad r_{y2} = 0.931 \quad r_{12} = 0.763$$

وبتطبيق صيغة معامل الارتباط الجزئي نحصل على:

$$r_{y2.1} = \frac{ry2 - (ry1)(r12)}{\sqrt{(1 - r^2y1)(1 - r^212)}}$$

$$r_{y2.1} = \frac{0.931 - (0.936)(0.763)}{\sqrt{(1 - 0.876)(1 - 0.582)}} = 0.86$$

وباختبار معنوية حجم معامل الارتباط الجزئي الذي مقداره 0.86 نستدل على مدى معنوية اضافة المتغير x_2 الى معادلة الانحدار ان كان سيؤدي الى مساهمة في تفسير المتغير التابع y . ويمكن الاستعانة بالجدول الإحصائية لمعامل ارتباط

Pearson المبين في الملحق لاختبار معنوية حجم معامل الارتباط، أو استخدام صيغة t التالية ومقارنتها مع قيمة t الجدولية عند مستوى معنوية α وكما يلي:

$$t = r_{y2.1} \sqrt{\frac{n - k - 1}{1 - r^2 y_{2.1}}}$$

$$= (0.86) (2.773) = 2.385$$

وبما قيمة t المحتسبة هي اقل من القيمة الجدولية $t_{0.025, 2} = 4.303$ نستدل على عدم معنوية معامل الارتباط الجزئي $r_{y2.1}$ رغم ان حجمه يبدو للوهلة الاولى كبيراً نسبياً، ويعود سبب ذلك الى صغر حجم العينة $n = 6$.

4. معامل ارتباط الرقب Rank Correlation Coefficient

ويدعى ايضا بمعامل ارتباط سبيرمان Spearman Correlation Coefficient الذي يستخدم مع البيانات غير الرقمية القابلة للترتيب التصاعدي والتنازلي، وهو متوفر ايضا وكما تطرقنا لذلك في برنامج SPSS، ويعود الى فصيلة الاحصاءات غير المعملية، ويرمز له r_s وصيغته هي:

$$r_s = 1 - \frac{6(\sum d_i^2)}{n(n^2 - 1)}$$

حيث ان d هي الفرق بين رتبة مشاهدة ما حسب المتغير الاول ورتبتها حسب المتغير الثاني، وعندما تكون هناك عدة مشاهدات بنفس الرتبة يعتبر الوسط الحسابي هو رتبة كل واحدة من تلك المشاهدات عند ترتيبها تصاعدياً.

مثال (4-5): في تقييم عينة تتكون من 11 طالب في لعبتي كرة الطائرة وكرة السلة، كانت نتائج التقييم كما في الجدول التالي، والمطلوب ايجاد العلاقة بين اداء الطالب في اللعبتين باستخدام معامل ارتباط الرتب.

| تسلسل الطالب | تقييم لعبة الطائرة x1 | تقييم لعبة السلة x2 |
|--------------|-----------------------|---------------------|
| 1 | جيد | ضعيف جدا |
| 2 | ضعيف | ممتاز |
| 3 | مقبول | ممتاز |
| 4 | جيد | جيد |
| 5 | ممتاز | مقبول |
| 6 | مقبول | جيد جدا |
| 7 | ضعيف جدا | مقبول |
| 8 | جيد جدا | ضعيف |
| 9 | ممتاز | جيد |
| 10 | ضعيف | ضعيف جدا |
| 11 | جيد جدا | مقبول |

الحل (4-5):

باعطاء رتبة كل طالب حسب مستوى التقييم في كلا اللعبتين يكون لدينا :

| التسلسل | x1 | x2 | di | d_i^2 |
|---------|------|------|----------------|------------------|
| 1 | 6.5 | 1.5 | 5 | 25 |
| 2 | 2.5 | 10.5 | -8 | 64 |
| 3 | 4.5 | 10.5 | -6 | 36 |
| 4 | 6.5 | 7.5 | -1 | 1 |
| 5 | 10.5 | 5 | 5.5 | 30.25 |
| 6 | 4.5 | 9 | -4.5 | 20.25 |
| 7 | 1 | 5 | -4 | 16 |
| 8 | 8.5 | 3 | 5.5 | 30.25 |
| 9 | 10.5 | 7.5 | 3 | 9 |
| 10 | 2.5 | 1.5 | 1 | 1 |
| 11 | 8.5 | 5 | 3.5 | 12.25 |
| | | | $\sum d = 0.0$ | $\sum d^2 = 245$ |

وبتطبيق صيغة معامل ارتباط الرتب التالية نحصل على :

$$r_s = 1 - \frac{6(\sum d_i^2)}{n(n^2 - 1)}$$

$$r_s = 1 - \frac{6(245)}{11(120)} = -0.114$$

اي ان العلاقة بين نتيجتي التقييم سالبة وضعيفة .

5. معامل الاقتران Coefficient of Association

ويستخدم في الحالات التي تكون فيها بيانات كلا المتغيرين يتكون من مستويين، وان احدهما او كلاهما غير قابلة للترتيب التصاعدي او التنازلي، ونرمز لمعامل الاقتران r_a . فاذا كانت a, b هي مستويات او حالات المتغير x و $1, 2$ هي مستويات المتغير y و n هي عدد التكرارات، اي :

| المتغير y | | المتغير X |
|-----------|----------|-----------|
| 2 | 1 | |
| n_{a2} | n_{a1} | a |
| n_{b2} | n_{b1} | b |

فان صيغة حساب معامل ارتباط الاقتران هي :

$$r_a = \frac{n_{a1}n_{b2} - n_{a2}n_{b1}}{n_{a1}n_{b2} + n_{a2}n_{b1}}$$

مثال (5-5): المطلوب ايجاد معامل الاقتران r_a بين ظاهرتي التدخين والمستوى التعليمي لعينة من الاشخاص حجمها $n = 120$ كما مبين في الجدول التالي:

| متغير الحالة التعليمية | | متغير حالة التدخين |
|------------------------|-----|--------------------|
| غير امي | امي | |
| 35 | 30 | يدخن |
| 15 | 40 | لا يدخن |

الحل (5-5):

بتطبيق صيغة معامل الاقتران اعلاه نحصل على:

$$r_a = \frac{(35)(40) - (30)(15)}{(35)(40) + (30)(15)} = 0.513$$

وفقا لحجم العينة الكبير نسبيا، فان العلاقة قوية وموجبة، اي ان نسبة التدخين تزداد بزيادة نسبة المتعلمين .

6. معامل التوافق Coefficient of Contingency

ويستهدف قياس الارتباط بين متغيرين احدهما او كلاهما ينقسم الى اكثر من حالتين (مستويين)، ويعتمد على استخدام مربعات كاي Chi Square، وصيغة احتسابه تاخذ الشكل التالي:

$$r_c = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

حيث إن :

$$\chi^2 = n \left\{ \frac{n^2_{11}}{nr_1nc_1} + \frac{n^2_{12}}{nr_1nc_2} + \dots + \frac{n^2_{rc}}{nr_1nc} \right\} - n$$

وترمز n^2_{11} الى مربع قيمة اول خلية واقعة في السطر الاول r_1 والعمود الاول c_1 تليها الخلية الثانية الواقعة في السطر الاول من العمود الثاني وهكذا لغاية آخر خلية تقع في آخر عمود وآخر سطر، بينما ترمز nr_1nc_1 الى حاصل ضرب مجموع العمود الاول في مجموع السطر الاول وهكذا .

مثال (5-6): المطلوب إيجاد معامل الارتباط التوافقي r_c بين متغيري المهنة الذي يشمل ثلاثة أنواع من المهن ومتغير التدخين المصنف الى حالتين.

| متغير التدخين | متغير المهنة | | | المجموع |
|---------------|--------------|----|----|---------|
| | a | b | c | |
| يدخن | 30 | 80 | 20 | 130 |
| لا يدخن | 25 | 15 | 30 | 70 |
| المجموع | 55 | 95 | 50 | 200 |

الحل (5-6): لدينا:

$$200 - \chi^2 = 200 \left\{ \frac{(30)^2}{(55)(130)} + \frac{(80)^2}{(95)(130)} + \frac{(20)^2}{(50)(130)} + \frac{(25)^2}{(55)(70)} + \frac{(15)^2}{(95)(70)} + \frac{(30)^2}{(35)(70)} \right\} - 200 = 31.76$$

وبالتعويض بالصيغة r_c نحصل على :

$$r_c = \sqrt{\frac{X^2}{X^2 + n}}$$

$$r_c = \sqrt{\frac{31.76}{31.76 + 200}} = 0.37$$

وعند الأخذ بنظر الاعتبار حجم العينة الكبير $n = 200$ فإن معامل الارتباط يشير الى علاقة قوية بين مهنة الشخص وحالة التدخين، مع التنويه ايضا الى زيادة الاعمدة والصفوف والذي من شأنه ان يزيد من معنوية معامل الارتباط ايضا .

تقارین الفصل الخامس

تمرین (1-5): الجدول التالي يضم بيانات لعينة من الموظفين تخص الاعمار (بالسنين) x_1 والخبرة الوظيفية (بالسنين) x_2 ومعدل الراتب الشهري (بالدينار) y ، والمطلوب:

1. استخدام برنامج SPSS لإيجاد معامل ارتباط Pearson مع توضيح معنوية واتجاه العلاقة .
2. إيجاد معامل الارتباط المتعدد R يدويا، وبيان مدى معنوية العلاقة .
3. إيجاد معامل الارتباط البسيط بين y و x_2 وبيان مدى معنوية واتجاه العلاقة.

| التسلسل | الراتب الشهري y | العمر x_1 | سنوات الخبرة x_2 |
|---------|-------------------|-------------|--------------------|
| 1 | 180 | 20 | 2 |
| 2 | 290 | 29 | 6 |
| 3 | 194 | 33 | 13 |
| 4 | 280 | 38 | 12 |
| 5 | 212 | 44 | 15 |
| 6 | 314 | 47 | 19 |
| 7 | 320 | 53 | 18 |
| 8 | 290 | 51 | 16 |

تمرین (2-5): تم الاستفسار من ربتي بيت عن رأيهن بعشرة انواع من مسحوق الغسيل، وكانت الاجابة كما هو مبين في الجدول التالي. والمطلوب :

1. تحويل البيانات النوعية الى كمية واستخراج معامل ارتباط Spearman باستخدام برنامج SPSS .
2. إيجاد معامل ارتباط الرتب يدويا .

| نوع المسحوق | رأي ربة البيت الأولى | رأي ربة البيت الثانية |
|-------------|----------------------|-----------------------|
| A | رديء | متوسط |
| B | رديء جداً | رديء |
| C | رديء | رديء جداً |
| D | جيد | متوسط |
| E | جيد | جيد جداً |
| F | ممتاز | جيد جداً |
| G | جيد جداً | جيد |
| H | متوسط | جيد |
| I | ممتاز جداً | ممتاز |

تمرين (3-5): استخدم الجدول التالي وعند $\alpha = 0.05$ لبيان ان كانت هناك علاقة ra بين مستوى الذكاء (وفقا لاختبار محدد) للبانعين في احد المخازن وبين حجم المبيعات.

| حجم المبيعات | مستوى الذكاء | | |
|--------------|--------------|-------|---------------|
| | أقل من متوسط | متوسط | أكثر من متوسط |
| قليل | 18 | 28 | 14 |
| متوسط | 37 | 63 | 30 |
| عالي | 15 | 29 | 16 |

تمرين (4-5): قام طبيبان نفسيان A.B بمقابلة 64 مريضا وسجلا فيما اذا كان المريض يعاني من انفصام بالشخصية ام لا وحصلا على التصنيف التالي، فهل هناك توافق في آراء الطبيبين في تشخيص المرض .

| الطبيب A | الطبيب B | |
|-----------------|-------------|-----------------|
| | المرض موجود | المرض غير موجود |
| المرض موجود | 21 | 8 |
| المرض غير موجود | 20 | 15 |



التحليل باستخدام الطرق متعددة المتغيرات

MULTIVARIATE TECHNIQUES

1.6. تحليل الانحدار Regression Analysis

1. مقدمة

يبحث الانحدار في العلاقة بين مجموعة المتغيرات المستقلة X_i والمتغير التابع Y_i من خلال بناء معادلة تستخدم للتنبؤ أو التفسير أو التقدير أو للتحكم والسيطرة. وبواسطة عملية التحليل يمكننا معرفة تأثير مجموعة المتغيرات المستقلة، وكذلك تأثير كل منها بصورة منفردة على المتغير التابع. والشكل العام لمعادلة الانحدار هو:

$$Y = \alpha + \beta X + e_i$$

حيث إن α تشير الى المعامل الثابت، و β معاملات (ميل) الانحدار، و e_i الخطأ العشوائي.

وعند بناء المعادلة التي تعتمد بيانات العينة التي من غير المتوقع ان تقع البيانات على خط الانحدار تماما، يصبح شكل المعادلة:

$$y = a + b_i x_i + e_i$$

ويتم تقدير ميل الانحدار غير العلوم باستخدام طريقة المربعات الصغرى التي تعتمد تقليل مجموع مربعات انحرافات القيم الحقيقية عن القيم التقديرية. ويخضع تحليل الانحدار لمجموعة فرضيات بالنسبة للمتغير العشوائي e_i ، ويتطلب التحقق منها قبل قرار قبول النموذج بصيغته النهائية، ومن أهمها هو: ان المتغيرات المستقلة والتابعة موزعة توزيعا طبيعيا normality؛ ومن أن تأثير المتغيرات يكون خطيا linearity؛ وعدم وجود علاقات معنوية او شبه تامة متداخلة multicollinearity بين طاقم المتغيرات المستقلة التي يتضمنها النموذج في حالة البيانات المقطعية و autocollinearity في حالة السلاسل الزمنية؛ ومن أن تباين الخطأ العشوائي ثابت؛ وان القيمة المتوقعة للخطأ العشوائي تساوي صفراً. وتعتبر الطريقة البيانية من اهم الطرق و ابسطها للتحقق من صحة هذه الفرضيات (للتفصيل يمكن الاستعانة ب "الاحصاء للعلوم الادارية والتطبيقية" للمؤلف). وهناك عدة معايير احصائية ومنطقية

يتم اعتمادها لاختبار معنوية النموذج والمتغيرات التي يتضمنها، بالإضافة الى استخدام البواقي residuals لاختبار معنوية نتائج تنبؤ النموذج، وسيتم التطرق لهذه المعايير بصورة مجملة عند تناول تفسير نتائج التحليل في الفقرة التالية، في حين يمكن الوقوف على صيغها النظرية في فقرة الطريقة اليدوية التي سيرد ذكرها لاحقاً.

2. استخدام الحاسوب مع برنامج SPSS

اولاً: اجراءات مدخلات تحليل الانحدار

بعد الدخول الى برنامج SPSS واختيار الامر الرئيسي Analyze يتم التأشير على الامر الفرعي Regression وستبدو لنا انواع مجالات الانحدار التي يمكن استخدامها في عملية التحليل وهي الخطية Linear ؛ ومجموعة Binary Logistic التي يأخذ متغيرها التابع قيمة احتمالية تقع بين الصفر والواحد ؛ ومجموعة غير الخطية Nonlinear التي تكون شكل العلاقة بين متغيرها التابع والمتغيرات المستقلة غير مستقيمة.

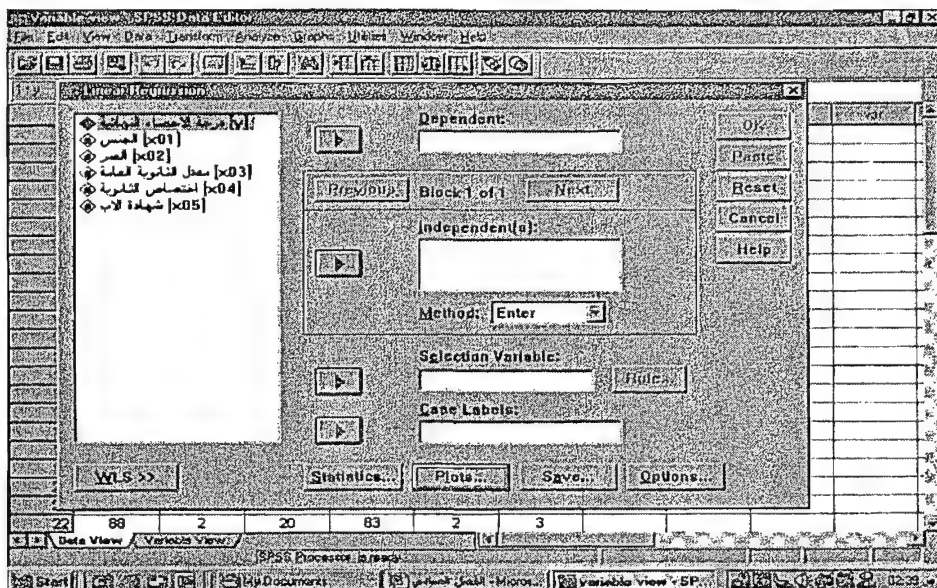
وستتابع تحليل النموذج الخطي باعتباره الأكثر استخداماً وأهمية، ولأن اجراءات استخدام الانواع الأخرى للانحدار في البرنامج متماثلة، بالإضافة لامكانية تحويل غير الخطية الى خطية من خلال اعادة صياغة المتغيرات، مستخدمين بيانات المثال (3-1) الذي يشمل عينة تتكون من 31 طالباً، لدراسة العوامل المؤثرة على اداء الطالب في مادة الاحصاء، حيث تمثل العلامات النهائية لمادة الاحصاء المتغير التابع Dependent Variable و 5 متغيرات مستقلة هي: الجنس (X1) و العمر (X2) ومعدل الثانوية العامة (X3) واختصاص الدراسة في الثانوية (X4) ومستوى التحصيل الدراسي للاب (X5).

فعند الكبس على خيار Linear سيظهر مربع الحوار المبين في الشكل رقم (6-1) وعليه تتضح مواقع كل من:

- ادخال المتغير التابع y والمتغيرات المستقلة Xi التي يمكن اختيارها من قائمة المتغيرات الموجودة على الجانب الايمن من مربع الحوار .
- طريقة التحليل Method المرغوب استخدامها، ومن اهمها طريقة Stepwise (الخطوات) التي بموجبها يتم اولا ادخال المتغير الذي يتصف باعلى معنوية في علاقته مع المتغير التابع، يليه المتغير المستقل الثاني الذي يلي الاول من ناحية المعنوية وهكذا ، والطريقة تتيح متابعة التغيرات التي تطرأ على النموذج عند اضافة كل متغير معنوي جديد، وتعتبر طريقة الخطوات من اكثر الطرق استخداما وشيوعا لما توفره من معلومات للباحث في كل خطوة جديدة من جهة، ولانها تحتاج لوقت اقل مما تحتاجه الطرق الاخرى في عملية التحليل؛ اما الطرق الاخرى فهي طريقة Enter التي استخدامها يعني ادخال كافة المتغيرات المستقلة في النموذج (المعادلة) ليقوم الباحث بتفحص معايير كل منها واختيار ما يراه مناسباً، او الابقاء عليها جميعا اذا كان النموذج يستهدف تفسير او وصف الظاهرة تحت الدراسة، خاصة ان كانت قائمة المتغيرات المرشحة للتحليل قد جاءت وفق خبرة سابقة عن تأثير كل منها على المتغير التابع، الا أن هذه الطريقة غير مناسبة بصورة كبيرة في حالة كان الهدف من بناء النموذج هو التنبؤ او بناء التوقعات المستقبلية التي يفضل معها ان يكون النموذج باقل عدد من المتغيرات اقتصادا في الكلفة ؛ وهناك ايضا طريقة Backward التي تتناول جميع المتغيرات ومن ثم تبدا باستبعاد المتغيرات مبدئاً من الاكثر غير معنوية ومن ثم الذي يليه من ناحية عدم المعنوية وهكذا، ولغاية التوقف عند المتغيرات التي تستوفي لدرجة المعنوية المقررة، بينما تقوم طريقة Forward بادخال كافة المتغيرات ايضا الا انها تبدأ باختيار المتغير الاكثر معنوية اولا والاستمرار على هذا المنوال والتوقف عند عدم استيفاء المتغير اللاحق لدرجة المعنوية المقررة .
- كما ويشتمل مربع الحوار ايضا تحديد المتغير المستقل Selection Variable المستهدف الابقاء عليه ضمن طاقم المتغيرات التي سيتضمنها النموذج، وذلك لاهميته المنطقية للظاهرة المدروسة من وجهة نظر الباحث .

- كذلك موقع التأثير على اظهار اسماء المتغيرات Labels بجانب رموزها ان كانت هناك رغبة او حاجة لذلك .
- بالإضافة الى توفر الايقونات المتعلقة باختيار مربعات حوار المعايير الاحصائية Statistics والتي توفر المعايير المتعلقة بالمعاملات ومعايير قياس معنوية النموذج ؛ واخرى تتعلق بالبقاوي (Residuals)، وايقونة Option التي توفر خيارات درجات المعنوية التي عندها يتم ادخال المتغير للتحليل والدرجة التي عندها يتم استبعاده وقيم Durban-Watson اذا كانت المشاهدات هي عبارة عن سلسلة زمنية، وايقونة الرسوم Plots للحصول على الشكل البياني لطبيعة العلاقة التي يظهر عليها المتغير التابع مع كل من المتغيرات المستقلة، اي ان لكل من الايقونات خيارات متعددة يمكن توظيفها لتكون ضمن المخرجات.

شكل بياني رقم (16) سريع حوار الانحدار الخطي



وعقب الانتهاء من العمل مع مربعات الحوار الملحقة بمربع الحوار الرئيسي يتم الكبس على ايقونة Ok للحصول على المخرجات (Output). وباخضاع ملف البيانات المتعلق مثالنا (3-1) المشار اليه للتحليل نحصل على المخرجات في الجدول (6-1) التالية:

جدول رقم (16)
مخرجات تحليل الانحدار Regression للمثال (13)

| | Mean | Std.Deviation | N |
|-----------------------|-------|---------------|----|
| درجة الإحصاء النهائية | 66.26 | 14.46 | 31 |
| الجنس | 1.45 | .51 | 31 |
| العمر | 21.55 | 4.37 | 31 |
| معدل الثانوية العامة | 65.55 | 7.68 | 31 |
| اختصاص الثانوية | 1.45 | .51 | 31 |
| شهادة الأب | 3.35 | .84 | 31 |

Correlations

| | شهادة الأب | اختصاص الثانوية | معدل الثانوية العامة | العمر | الجنس | درجة الإحصاء النهائية |
|---------------------|------------|-----------------|----------------------|-------|-------|-----------------------|
| Pearson Correlation | .242 | .348 | .601 | .233 | -.080 | درجة الإحصاء النهائية |
| | -.469 | -.172 | -.014 | -.040 | 1.000 | الجنس |
| | .082 | .035 | .115 | 1.000 | -.040 | العمر |
| | .460 | .578 | 1.000 | .115 | -.014 | معدل الثانوية العامة |
| | .317 | 1.000 | .578 | .035 | -.172 | اختصاص الثانوية |
| | 1.000 | .317 | .460 | .082 | -.469 | شهادة الأب |
| Sig.(2-tailed) | .094 | .027 | .000 | .104 | .334 | درجة الإحصاء النهائية |
| | .004 | .177 | .469 | .415 | 0 | الجنس |
| | .331 | .426 | .269 | 0 | .415 | العمر |
| | .005 | .000 | 0 | .269 | .469 | معدل الثانوية العامة |
| | .041 | 0 | .000 | .426 | .177 | اختصاص الثانوية |
| | 0 | .041 | .005 | .331 | .004 | شهادة الأب |
| N | 31 | 31 | 31 | 31 | 31 | درجة الإحصاء النهائية |
| | 31 | 31 | 31 | 31 | 31 | الجنس |
| | 31 | 31 | 31 | 31 | 31 | العمر |
| | 31 | 31 | 31 | 31 | 31 | معدل الثانوية العامة |
| | 31 | 31 | 31 | 31 | 31 | اختصاص الثانوية |
| | 31 | 31 | 31 | 31 | 31 | شهادة الأب |

Variables Entered/ Removed^a

| Model | Variables Entered | Variables Removed | Method |
|-------|----------------------|-------------------|--|
| 1 | معدل الثانوية العامة | | Stepwise (Criteria: Probability-of-F-to-enter ≤ .050, Probability-of-F-to-remove ≥ .050) |

Model Summary^a

| Model | R | R Square | Adjusted R Square | Std. Error of the Estimate | R Square Change | F Change | df1 | df2 | Sig. F Change |
|-------|-------------------|----------|-------------------|----------------------------|-----------------|----------|-----|-----|---------------|
| 1 | .601 ^b | .361 | .339 | 11.75 | .361 | 16.415 | 1 | 29 | .000 |

^a - Dependent Variable: درجة الإحصاء النهائية

^b - Predictors: (Constant): معدل الثانوية العامة

ANOVA^a

| Model | | Sum of Squares | Df | Mean Square | F | Sig. |
|-------|------------|----------------|----|-------------|--------|-------------------|
| 1 | Regression | 2266.209 | 1 | 2266.209 | 16.415 | .000 ^b |
| | Residual | 4003.727 | 29 | 138.060 | | |
| | Total | 6269.935 | 30 | | | |

Coefficients^a

| Model | | Unstandardized Coefficients | | Standardized Coefficients | | Sig. | Correlations | | |
|-------|----------------------|-----------------------------|------------|---------------------------|-------|------|--------------|---------|------|
| | | B | Std. Error | Beta | t | | Zero-order | Partial | Part |
| 1 | (Constant) | -7.918 | 18.430 | | -.430 | .671 | | | |
| | معدل الثانوية العامة | 1.132 | .279 | .601 | 4.052 | .000 | .601 | .601 | .601 |

Excluded Variables^a

| Model | | Beat In | T | Sig. | Partial Correlation | Collinearity Statistics |
|-------|-----------------|--------------------|-------|------|---------------------|-------------------------|
| | | | | | | Tolerance |
| 1 | الجنس | -.072 ^b | -.476 | .638 | -.090 | 1.000 |
| | العمر | .166 ^b | 1.114 | .275 | .206 | .987 |
| | اختصاص الثانوية | .001 ^b | .008 | .994 | .001 | .666 |
| | شهادة الأب | -.044 ^b | -.257 | .799 | -.049 | .788 |

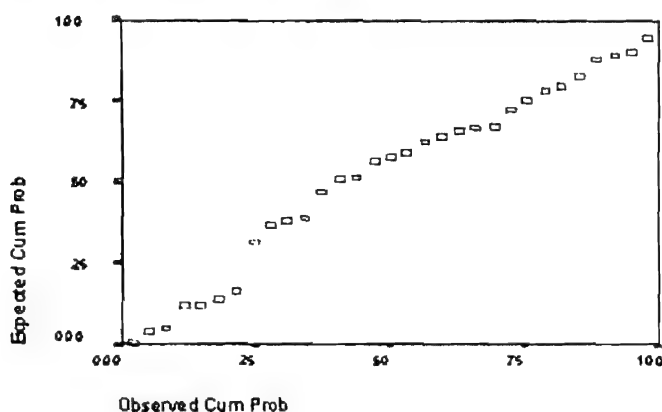
^a - Dependent Variable: درجة الإحصاء النهائية

^b - Predictors: (Constant): معدل الثانوية العامة

Residuals Statistics^a

| | Minimum | Maximum | Mean | Std. Deviation | N |
|---------------------|---------|---------|-----------|----------------|----|
| Predicted Value | 54.32 | 86.01 | 66.26 | 8.69 | 31 |
| Residual | -31.30 | 18.57 | -9.17E-15 | 11.55 | 31 |
| Std.Predicted Value | -1.373 | 2.272 | .000 | 1.000 | 31 |
| Std. Residual | -2.663 | 1.581 | .000 | .983 | 31 |

Normal P-P Plot of Regression Standardized Residual



ثانياً: تفسير مخرجات تحليل الانحدار باستخدام برنامج

SPSS

من جدول المخرجات (6-1) اعلاه نجد ان نتائج تحليل الانحدار التي جاءت ضمن المخرجات ووفق ما تم اختياره من طرق وعمليات تحليلية ومعايير لمقياس معنوية نتائج التحليل من بين ما هو متاح في البرنامج تشمل ما يلي :

^a - Dependent Variable: درجة الإحصاء النهائية

- قائمة بأسماء المتغيرات التي تم اخضاعها لعملية التحليل مع مقاييس الوسط الحسابي والانحراف المعياري وعدد المشاهدات لكل من هذه المتغيرات .
- مصفوفة الارتباط لكافة المتغيرات، والتي اعتمدت عليها عملية التحليل وتوضح معامل الارتباط بين المتغير التابع وكل من المتغيرات المستقلة، وكذلك درجة العلاقة فيما بين المتغيرات المستقلة ذاتها، مع اتجاه هذه العلاقات (الإشارة) .
- اسم المتغير او المتغيرات المستقلة التي تم ادخالها في النموذج ذات التأثير المعنوي على الظاهرة (المتغير التابع) وفقا لمعيار المعنوية المقرر مسبقا.
- مقاييس معنوية النموذج الذي تم تطويره وهي: R , R^2 , F - ratio test , وجميعها وكما هو مبين في اعلاه ذات معنوية عالية $\text{Significance} = 0.000$ ، كما وان اشارة المتغير المستقل الداخل في النموذج قد جاءت باشارة موجبة، وهذا يعني انه كلما كانت معدلات الطلبة في الثانوية العامة مرتفعة يزداد مستوى ادائهم في مادة الاحصاء. واصبح شكل نموذج الانحدار كالآتي:

$$y = -7.918 + 1.132 \times x_3$$

$$t = - 0.430 \quad 4.052$$

$$R = 0.601$$

$$R^2 = 0.361$$

$$F = 16.415 \quad , \quad \text{Sig.} = 0.000$$

- تحليل التباين ANOVA للوقوف على درجة التقارب وتحليل اسباب الاختلاف بين القيم الحقيقية والقيم التي تم الحصول عليها باستخدام النموذج المطور .
- مقاييس معنوية كل من المعامل الثابت Constant ومعاملات انحدار المتغيرات المستقلة التي تضمنها النموذج باستخدام المعايير: t -test ومعامل

- الارتباط الجزئي Partial Correlation Coefficient لكل من المتغيرات الداخلة في النموذج، وجميع هذه المعايير جاءت عالية المعنوية (0.000) .
- قائمة بالمتغيرات التي تم استبعادها ودرجة معنوية كل منها والتي عادة ما تكون منخفضة وفقاً لدرجة المعنوية المقررة مما أدى إلى استبعادها.
 - المقاييس المتعلقة بالبواقي المعيارية Standardized وموضحة في الشكل البياني والذي منه يستدل على الكفاءة العالية للنموذج المطور من خلال ملاحظة التقارب الشديد للقيم المستخرجة بواسطة النموذج المطور من الخط المستقيم (النموذجي).

2.6 تحليل المركبات Principal Component Analysis

1. مقدمة

ويعتبر تحليل المركبات (بضم الميم) أهم فصيلة في تحليل العوامل Factor Analysis لامكانية استخدامه مع البيانات سواء اكانت موزعة طبيعياً ام لا. وهو عبارة عن اداة وصفية تستطيع تصنيف اعداد كبيرة من المتغيرات الى عدد محدود من المركبات (العوامل) اعتماداً على العلاقات التي تربط كل مجموعة من المتغيرات فيما بينها، وهو بذلك يستطيع تقليص عدد كبير من المتغيرات من دون ان يؤدي ذلك الى فقدان جوهرى في نسبة التباين التي يتم تفسيرها، وهو اجراء يساعد على التخلص من مشكلة العلاقات المتداخلة Multicollinearity التي تواجه المتغيرات المستقلة. بكلمة اخرى تكون المتغيرات التي تضمها كل مركبة مترابطة بينها، بينما تكون العلاقة بين المركبات غير مترابطة. وعند المخرجات يأتي تسلسل العوامل وفقاً لحجم التباين الذي يستطيع كل مركب تفسيره بواسطة المتغيرات التي يتضمنها، فالمركب الاول هو الذي يفسر اعلى نسبة من التباين ويليه المركب الثاني وهكذا. ولتحقيق خاصية عدم الترابط بين المركبات يجب اختيار طريقة Orthogonal، اما في حالة البحث عن العلاقة بين المركبات فيتم استخدام طريقة Oblique Rotation وهو ما يطلق عليه بالتعامد. وعادة ما تستمر عملية التحليل ولغاية التفسير التام للتباين، او التوقف عند درجة المعنوية المقررة. مع

الإشارة هنا الى صعوبة انجاز هذا التحليل من دون استخدام الحاسوب لتعدد المصفوفات وسعة العمليات التحليلية المطلوبة لغاية الحصول على المخرجات، لذا سيقصر التطرق الى هذا النوع من التحليل المتقدم في حالة استخدام الحاسوب فقط. فلو رمزنا للمركبات الاساسية Cps كمتغيرات عشوائية غير مترابطة تضم مجموعة متغيرات X_1, X_2, \dots, X_p مترابطة وتشارك باتجاه خطي فستأخذ الصيغة التالية :

$$C_j = \sum_{i=1}^p a_i$$

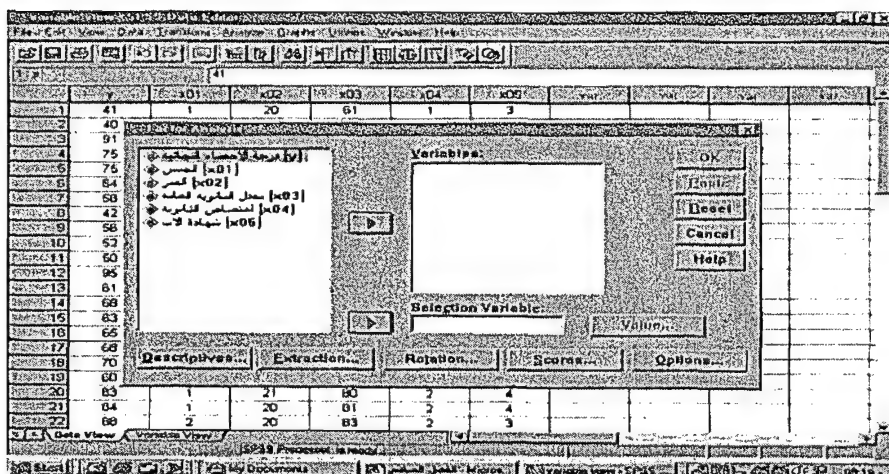
حيث ان:

$$j = 1, 2, \dots, p$$

2 اجراءات مدخلات تحليل المركبات

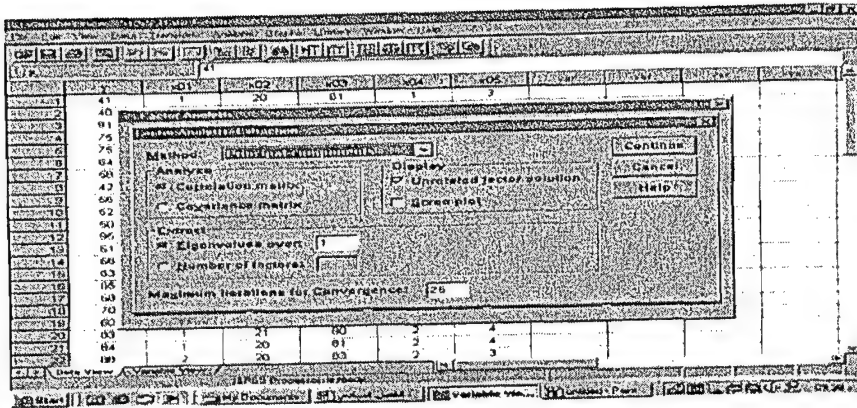
بعد الدخول على برنامج SPSS والكبس على الامر الرئيسي Analyze يتم اختيار الامر الفرعي Data Reduction والكبس على طريقة Factor Analysis فنحصل على مربع الحوار المبين في الشكل البياني رقم (2-6).

شكل بياني رقم (26)
مربع حوار تحليل المركبات



- وكما هو مبين على مربع الحوار تظهر مجموعة المتغيرات المزمع اخضاعها للتحليل على الجانب الايسر فيتم ادخالها في مربع الجانب الايمن، كما يوجد في الاسفل مجموعة ايقونات هي من اليسار الى اليمين تشمل العناوين التالية:
- **Descriptive**: وفي حالة استخدامها سيظهر مربع حوار ملحق يخص مصفوفة الارتباط بشأن تضمين المخرجات المعاملات ومستوى معنويتها، وكذلك الحل الاول Initial Solution .
 - **Extraction**: وفي هذا المربع الملحق وكما مبين في الشكل البياني رقم (3.6) هناك حقل يحمل عنوان Method وفيه يتم اختيار طريقة التحليل وهي Principal Component Analysis وفي حقل Analyze يمكن الخيار بين مصفوفة الارتباط او مصفوفة التباين المشترك Covariance Matrix، وهناك حقل Extract وفيه يمكن تحديد الحد الادنى للتباين الذاتي Eigenvalue وقد يكون من المناسب اختيار 1 كحد ادنى مثلا مع توفر خيار تحديد عدد العمليات التحليلية Maximum Iteration for Convergence .
 - **Rotation**: وفيها عرض خيارات استخدام طريقة Varimax (الاكثر استخداما) او غيرها من الطرق المتوفرة اوبدونها .
 - **Score**: وتخص خيار عرض قيم مصفوفة الارتباط.
 - **Option**: وفيها يمكن تحديد الحد الادنى لحجم معاملات الارتباط المطلوب في المخرجات.

شكل رقم (3.6) مربع الحوار للمقولة لاستخراج



وباكتمال البت في الخيارات المعروضة سواء في مربع الحوار الرئيسي او المربعات الملحقة يتم الكبس على ايقونة Ok للحصول على مخرجات عملية تحليل بيانات المثال (3-1) والمبين نموذجها في الجدول رقم (2-6)

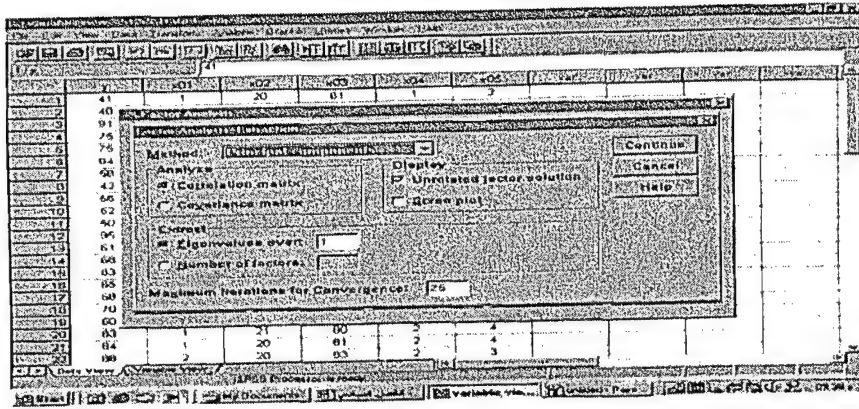
جدول رقم (2.6) نموذج مخرجات تحليل المركبات Principal Component Analysis

| شهادة الأب | اختصاص الثانوية | معدل الثانوية العامة | العمر | الجنس | درجة الإحصاء النهائية | |
|------------|-----------------|----------------------|-------|-------|-----------------------|-----------------------|
| .242 | .348 | .601 | .233 | -.080 | 1.000 | درجة الإحصاء النهائية |
| -.469 | -.172 | -.014 | -.040 | 1.000 | -.080 | الجنس |
| .082 | .035 | .115 | 1.000 | -.040 | .233 | العمر |
| .460 | .578 | 1.000 | .115 | -.014 | .601 | معدل الثانوية العامة |
| .317 | 1.000 | .578 | .035 | -.172 | .348 | اختصاص الثانوية |
| 1.000 | .317 | .460 | .082 | -.469 | .242 | شهادة الأب |
| .094 | .027 | .000 | .104 | .334 | 0 | درجة الإحصاء النهائية |
| .004 | .177 | .469 | .415 | 0 | .334 | الجنس |
| .331 | .426 | .269 | 0 | .415 | .104 | العمر |
| .005 | .000 | 0 | .269 | .469 | .000 | معدل الثانوية العامة |
| .041 | 0 | .000 | .426 | .177 | .027 | اختصاص الثانوية |
| 0 | .041 | .005 | .331 | .004 | .094 | شهادة الأب |

وكما هو مبين على مربع الحوار تظهر مجموعة المتغيرات المزمع اخضاعها للتحليل على الجانب الايسر فيتم ادخالها في مربع الجانب الايمن، كما يوجد في الاسفل مجموعة ايقونات هي من اليسار الى اليمين تشمل العناوين التالية:

- **Descriptive** : وفي حالة استخدامها سيظهر مربع حوار ملحق يخص مصفوفة الارتباط بشأن تضمين المخرجات المعاملات ومستوى معنويتها، وكذلك الحل الاول Initial Solution .
- **Extraction**: وفي هذا المربع الملحق وكما مبين في الشكل البياني رقم (3.6) هناك حقل يحمل عنوان Method وفيه يتم اختيار طريقة التحليل وهي Principal Component Analysis وفي حقل Analyze يمكن الخيار بين مصفوفة الارتباط او مصفوفة التباين المشترك Covariance Matrix، وهناك حقل Extract وفيه يمكن تحديد الحد الادنى للتباين الذاتي Eigenvalue وقد يكون من المناسب اختيار 1 كحد ادنى مثلاً مع توفر خيار تحديد عدد العمليات التحليلية Maximum Iteration for Convergence .
- **Rotation**: وفيها عرض خيارات استخدام طريقة Varimax (الاكثر استخداماً) او غيرها من الطرق المتوفرة اوبدونها .
- **Score**: وتخص خيار عرض قيم مصفوفة الارتباط.
- **Option**: وفيها يمكن تحديد الحد الادنى لحجم معاملات الارتباط المطلوب في المخرجات.

شكل رقم (36)
سريع الحوار الملحق لايقونة Extraction



وباكتمال البت في الخيارات المعروضة سواء في مربع الحوار الرئيسي او المربعات الملحقة يتم الكبس على ايقونة Ok للحصول على مخرجات عملية تحليل بيانات المثال (1-3) والمبين نموذجهما في الجدول رقم (2-6)

جدول رقم (26)
نموذج مخزومات تحليل المركبات Principal Component Analysis

| شهادة الأب | اختصاص ثانوية | معدل الثانوية العامة | العمر | الجنس | درجة الإحصاء النهائية | |
|------------|---------------|----------------------|-------|-------|-----------------------|---|
| .242 | .348 | .601 | .233 | -.080 | 1.000 | درجة الإحصاء النهائية Pearson Correlation |
| -.469 | -.172 | -.014 | -.040 | 1.000 | -.080 | الجنس |
| .082 | .035 | .115 | 1.000 | -.040 | .233 | العمر |
| .460 | .578 | 1.000 | .115 | -.014 | .601 | معدل الثانوية العامة |
| .317 | 1.000 | .578 | .035 | -.172 | .348 | اختصاص الثانوية |
| 1.000 | .317 | .460 | .082 | -.469 | .242 | شهادة الأب |
| .094 | .027 | .000 | .104 | .334 | 0 | درجة الإحصاء النهائية Sig.(2-tailed) |
| .004 | .177 | .469 | .415 | 0 | .334 | الجنس |
| .331 | .426 | .269 | 0 | .415 | .104 | العمر |
| .005 | .000 | 0 | .269 | .469 | .000 | معدل الثانوية العامة |
| .041 | 0 | .000 | .426 | .177 | .027 | اختصاص الثانوية |
| 0 | .041 | .005 | .331 | .004 | .094 | شهادة الأب |

Communalities

| | Initial | Extraction |
|-----------------------|---------|------------|
| درجة الإحصاء النهائية | 1.000 | .666 |
| الجنس | 1.000 | .843 |
| العمر | 1.000 | .930 |
| معدل الثانوية العامة | 1.000 | .841 |
| اختصاص الثانوية | 1.000 | .633 |
| شهادة الأب | 1.000 | .726 |

Extraction Method: Principal Component Analysis

Total Variance Explained

| Component | Initial Eigenvalues | | | Extraction Sums of Squared Loading | | | Rotation Sums of Squared Loadings | | |
|-----------|---------------------|---------------|-------------|------------------------------------|---------------|-------------|-----------------------------------|---------------|-------------|
| | Total | % of Variance | Cumulative% | Total | % of Variance | Cumulative% | Total | % of Variance | Cumulative% |
| 1 | 2.423 | 40.376 | 40.376 | 2.423 | 40.376 | 40.376 | 2.115 | 35.244 | 35.244 |
| 2 | 1.213 | 20.216 | 60.592 | 1.213 | 20.216 | 60.592 | 1.454 | 24.238 | 59.482 |
| 3 | 1.002 | 16.706 | 77.298 | 1.002 | 16.706 | 77.298 | 1.069 | 17.816 | 77.298 |
| 4 | .593 | 9.885 | 87.183 | | | | | | |
| 5 | .541 | 9.010 | 96.193 | | | | | | |
| 6 | .228 | 3.807 | 100.000 | | | | | | |

Extraction Method: Principal Component Analysis

Component Matrix^a

| | Component | | |
|-----------------------|-----------|-------|-------|
| | 1 | 2 | 3 |
| معدل الثانوية العامة | .843 | .301 | |
| اختصاص الثانوية | .722 | | -.323 |
| درجة الإحصاء النهائية | .713 | .380 | |
| شهادة الأب | .692 | -.496 | |
| الجنس | -.372 | .815 | |
| العمر | | | .897 |

Extraction Method: Principal Component Analysis

Rotated Component Matrix^b

| | Component | | |
|-----------------------|-----------|-------|-------|
| | 1 | 2 | 3 |
| معدل الثانوية العامة | .843 | .301 | |
| اختصاص الثانوية | .722 | | -.323 |
| درجة الإحصاء النهائية | .713 | .380 | |
| شهادة الأب | .692 | -.496 | |
| الجنس | -.372 | .815 | |
| العمر | | | .897 |

Extraction Method: Principal Component Analysis

Rotation Method: Varimax with Kaiser Normalization.

Component Transforamtion Matrix

| Component | 1 | 2 | 3 |
|-----------|------|------|-------|
| 1 | .843 | .301 | |
| 2 | .722 | | -.323 |
| 3 | .713 | .380 | |

Extraction Method: Principal Component Analysis

Rotation Method: Varimax with Kaiser Normalization.

^a – 3components extracted.

^b – Rotation converged in 5 iterations.

3. تفسير مخرجات تحليل المركبات

- وفقا لما تم تحديده من خيارات لاغراض التحليل فان النتائج التي تضمنها جدول المخرجات تشمل الآتي:
- مصفوفة الارتباط تتضمن حجم واسارة المعاملات، ومنها يتضح ان اقوى علاقة هي بين المتغيرين معدل الثانوية العامة ومستوى الاداء في مادة الاحصاء، وقد بلغ معامل الارتباط بينهما 0.60 وكما هو واضح جاء باشارة موجبة مما يعني ان الطالب الحاصل على معدل اعلى في الثانوية العامة يكون ادائه افضل في مادة الاحصاء .
 - ان نسبة التباين المشترك (Communality) المفسر بواسطة المتغيرات يتراوح بين 0.67 و 0.84.
 - ان مجموع التباين التراكمي المفسر للمركبات الثلاثة الاولى هو 0.773 ، وان كلاً من هذه المركبات الثلاثة استطاع تفسير التباين الذاتي Eigenvalue ما قيمته اكثر من 1.
 - ان اعلى معامل تحميل Loading في كل من المركبات الثلاثة التي آلت اليه عملية التحميل تعود على التوالي الى: متغير معدل الثانوية العامة (0.911)، ومتغير العمر (0.916)، ومتغير الجنس (0.962). وبذلك يمكن التعبير عن 6 متغيرات بثلاثة عوامل فقط. يمكن اشتقاق اسماء هذه العوامل من اسماء المتغيرات التي حققت اعلى تحميل في كل منها .

3.6 الطريقة اليدوية في تحليل الانحدار الخطي

1. مقدمة

كما هو معلوم فان شكل انتشار البيانات هو الذي يدلنا ان كان النموذج خطياً او غير خطي، حيث في الحالة الاخيرة يأخذ شكل منحني بدلا من الخط المستقيم الذي يكون عليه في حالة الخطي. ونعود الى المعادلة التقديرية التي سبق التطرق اليها في مقدمة هذا الفصل وهي:

$$y = a + b x_i$$

فكما نلاحظ نحن بحاجة الى ايجاد قيم كل من a و b باستخدام طريقة المربعات الصغرى (للزيادة في التفصيل يمكن الرجوع للمصدر السابق) ويتم ذلك باستخدام الصيغ التالية :

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}$$

$$a = \frac{\sum y - b \sum x}{n}$$

مثال (2-6): اخذت مؤشرات التطور الحاصل خلال آخر 5 سنوات ل 7 بلديات عن كل من متغير تطور عدد الإناث في التعليم الجامعي، ومتغير نسبة الزيادة في دخل الأسرة السنوي كما هو مبين في الجدول ادناه، والمطلوب تحديد معادلة الانحدار .

| البلدية | عدد الاناث في التعليم الجامعي (بالاف) y | نسبة زيادة معدل دخل الاسرة السنوي (%) x |
|---------|---|---|
| 1 | 10 | 5 |
| 2 | 11 | 9 |
| 3 | 18 | 11 |
| 4 | 17 | 13 |
| 5 | 15 | 14 |
| 6 | 21 | 18 |
| 7 | 24 | 21 |

الحل (2-6):

وفقا لمتطلبات الصيغ اعلاه لدينا:

$$\sum x^2 = 1357, \sum y = 116, \sum x = 91, \sum xy = 1660$$

نستخرج قيم كل من a و b باستخدام الصيغ اعلاه وكالاتي :

$$b = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2} = \frac{(7)(1660) - (116)(91)}{(7)(1357) - (8281)} = 0.874$$

$$a = \frac{\sum y - b \sum x}{n} = \frac{116 - (0.874)(91)}{7} = 5.21$$

وبذلك تكون معادلة الانحدار التقديرية هي:

$$y = 5.21 + 0.874 x$$

اما صيغ المعايير الاحصائية التي تستخدم لاختبار معنوية النموذج والمعاملات التي يتضمنها فهي :

اختبار t : لاختبار معنوية كل من معاملات الانحدار التي يتضمنها النموذج وصيغته :

$$t = b / Sb$$

حيث ان: b هو ميل الانحدار، Sb الخطأ المعياري لميل الانحدار
معامل التحديد R^2 : يوضح نسبة التباين التي يمكن تفسيرها بواسطة النموذج
وصيغة حسابه هي:

$$R^2 = \sum (y - \hat{y})^2 / \sum (y - \bar{y})^2$$

حيث ان: y قيم المتغير التابع، \bar{y} متوسط قيم المتغير التابع، \hat{y} القيم
التقديرية المستخرجة بواسطة المعادلة (النموذج).

اختبار F : و يستخدم لاختبار العلاقة بين طاقم المتغيرات المستقلة X_i
والمتغير التابع y وصيغة حسابه هي :

$$F = (R^2/k) / [(1-R^2) / (n-k-1)]$$

حيث ان: $n-k-1$ درجات الحرية، k عدد المتغيرات المستقلة الداخلة في
المعادلة.

2 استخدام نموذج الانحدار للتنبؤ

عقب تقييم نموذج الانحدار والتأكد من استيفائه للمعايير الاحصائية والمنطقية والفرضيات، يصبح بالإمكان استخدامه لأغراض التنبؤ، ويتمثل ذلك بتعويض القيم المطلوبة في X للحصول على قيم y . فبالنسبة لمعادلة الانحدار التي تم بناؤها لاداء الطلبة في مادة الاحصاء التي باستخدام برنامج SPSS وهي:

$$y = -7.918 + 1.132 x_3$$

وكنا بصدد التنبؤ بعلامة الاحصاء في الامتحان النهائي، نقوم بتعويض معدل لطالب في الثانوية العامة في x_3 ولنفترض كان 88 فنحصل على :

$$y = -7.918 + 1.132 (88) = 98.8$$

وهي علامة الاحصاء المتوقع الحصول عليها في الاحصاء اذا كان معدله في الثانوية العامة هو 88.

تمارين الفصل السادس

تمرين (6-1): مدير إحدى الشركات أراد بناء نموذج لتقييم الموظفين (y) العاملين لدى شركته وفقاً لمعدل إنجازيتهم (x) وذلك حسب المعلومات السابقة المتوفرة في الشركة فاختار عينة عشوائية تتكون من 10 موظفين وكانت المعلومات كما هي مبينة في الجدول التالي:

| الموظف | درجة التقييم (y) | مقدار الانجازية (x) |
|--------|----------------------|-------------------------|
| 1 | 75 | 70 |
| 2 | 64 | 71 |
| 3 | 92 | 92 |
| 4 | 80 | 80 |
| 5 | 76 | 48 |
| 6 | 58 | 64 |
| 7 | 96 | 90 |
| 8 | 89 | 75 |
| 9 | 98 | 86 |
| 10 | 76 | 58 |

والمطلوب:

1. بناء معادلة انحدار يدويا وباستخدام برنامج SPSS بتوظيف طريقة المربعات الصغرى .
2. إيجاد تقديرات لمعدلات تقييم الموظفين العشرة باعتماد نموذج الانحدار الذي يتم بناؤه .
3. التنبؤ بمعدل تقييم موظف مقار إنجازيته هي 95 .

- تمرين (6-2): أ. بين أهداف استخدام تحليل العوامل بطريقة المركبات .
- ب. وضح إجراءات استخدام برنامج SPSS لتحليل العوامل وفق طريقة المركبات Principal Component Analysis على أن تشمل المخرجات المركبات التي لا يقل تباينها الذاتي (Eigenvalue) عن 1 .
- ج. شرح كيفية اختيار أسماء المركبات التي تحصل عليها في المخرجات.



اختبار الفروض وتحليل التباين

HYPOTHESES TESTING AND ANALYSIS OF VARIANCE

1.7. مقدمة

- وهو احد المواضيع الرئيسية للاستدلال الاحصائي Inferential Statistics ويستهدف الوصول الى قرار القبول او الرفض بشأن :
- تقدير المعلمة المعتمدة على بيانات العينة المسحوبة من مجتمع المعلمة للتوصل الى درجة اعتمادية وثقة نتائج العينة .
 - اختبار الفروق بين النتائج الفعلية للعينة والنتائج الفرضية المتوقعة .
- ويمكن اجمال الاسس التي يقوم عليها اختبار الفروض بما يلي:

1. الفروض Hypotheses

الاولى وتسمى بفرضية العدم null hypothesis ويرمز لها H_0 وهي تتضمن الهدف المطلوب اختباره، ففي حالة قبولها يعني انها متوافقة مع الهدف، اي عدم وجود ما يدعو الى رفض النتائج. **والثانية** وتسمى بالفرضية البديلة alternative hypothesis ويرمز لها H_1 ، فعند رفض H_0 يعني قبول H_1 والعكس صحيح.

فمثلا اذا اردنا اختبار فرضية من ان متوسط وزن الطالب في الجامعة هو 62 كغم فان صيغة الفرضيات ستكون على الشكل التالي:

$$H_0: \mu = 62$$

$$H_1: \mu \neq 62$$

2. الخطأ من النوع الاول Type I error والخطأ من النوع

الثاني Type II error

عند رفض فرضية العدم H_0 ولكن كان يجب قبولها، لان عملية الرفض هو نتيجة خطأ في البيانات، عندها نقع في الخطأ من النوع الاول، وان احتمال الوقوع في مثل هذا الخطأ يرمز له α وتدعى بمستوى المعنوية (الدلالة). وكلما تقل قيمة α كلما قل احتمال الوقوع في الخطأ من النوع الاول .

اما الخطأ من النوع الثاني فيقع في حالة قبولنا لفرضية العدم H_0 بينما هذا القبول هو خطأ، وان احتمال الوقوع في هذا النوع من الخطأ يرمز له β .

3. اختبار من جانب واحد tail واختبار من جانبيين II tails

ويقصد به ان الانحراف عن فرضية العدم هو باتجاه واحد او انها موزعة على جانبيين. وهذا يعتمد على صيغة فرضية العدم، فاذا كانت الاشارة هي \geq (اكبر من او يساوي) او \leq (اقل من او يساوي) اي:

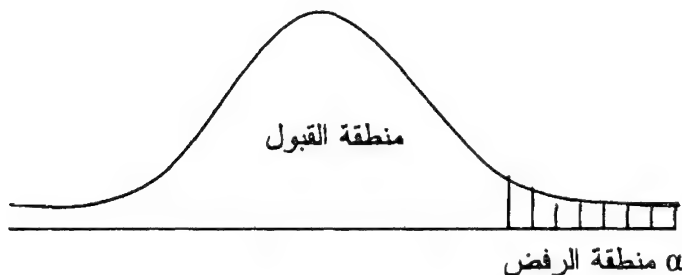
$$H_0 : \mu \geq 62$$

$$H_1 : \mu \leq 62$$

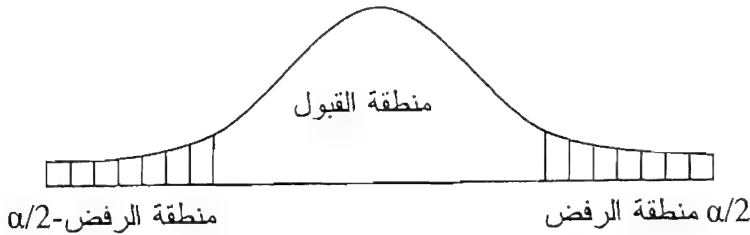
فهذا يعني بان الاختبار من جانب واحد، لانه في حالة رفض الفرضية فمن المتوقع حصرا بان الفرضية البديلة سيكون معلوما اتجاها كما هو مبين من الشكل البياني (1-7). اما في حالة ان تكون فرضية العدم مع اشارة =، فهذا يعني عدم معلومية الاتجاه الذي ستكون عليه في حالة رفضها فقد تكون اقل من او اكبر من، وبذلك ستتوزع على جانبيين كما هو مبين في الشكل البياني رقم (2-7).

شكل بياني رقم (1-7)

موضع منطقة الرفض في اختبار من جانب واحد



شكل بياني رقم (27)
يوضح منطقة الرفض في حالة اختبار من جانبيين



2.7. استخدام الحاسوب مع برنامج SPSS

1. الاختبار الاحادي One Sample T- Test

اولا: المفهوم والمدخلات

ويقصد به اختبار X (او متوسط العينة) مع متوسط المجتمع μ ، للتوصل ان كان هناك فرق جوهري بينهما، وعلى افتراض تساوي التباين، كما هو الحال لو كنا بصدد اختبار ادعاء احد فروع البنك مع بيانات البنك الرئيسي مثلا، او بصدد اختبار عينة من منتجات شركة صناعية للتأكد من مطابقتها لخصائص ومواصفات انتاج الشركة، اي ان العينة مسحوبة من ذات الشركة .

مثال (7-1): مصنع لانتاج معدات الرياضية ادعى بانه استطاع صناعة مضرب للتنس

بمقاومة متوسطها $\mu = 6.5$ كغم وبانحراف معياري $\sigma = 0.45$ والمطلوب

اختبار ادعاء المصنع مع نتائج عينة حجمها $n = 62$ مضرب والمبينة

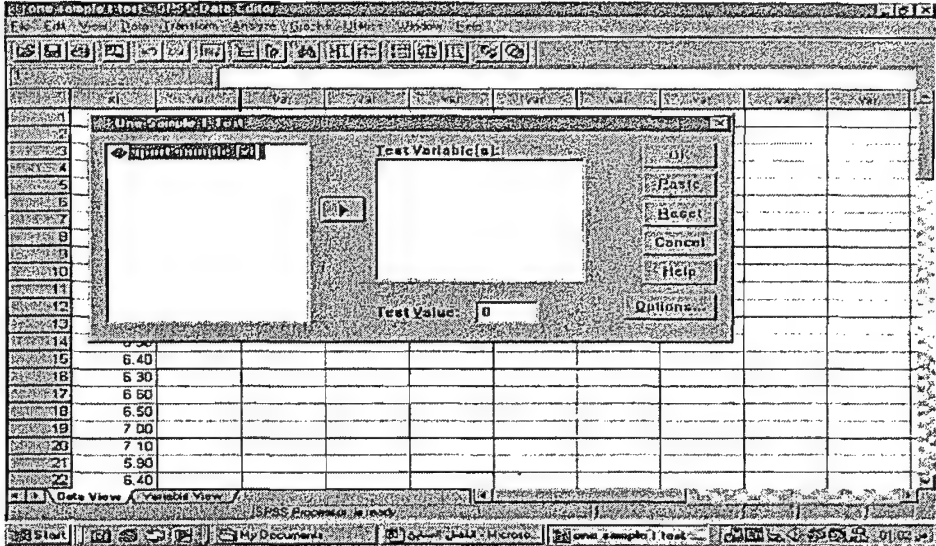
قيمتها في ادناه :

6.4، 6.3، 6.0، 6.5، 6.5، 7.0، 7.1، 6.5، 5.9، 6.4، 6.6، 6.7، 6.7،
6.5، 6.4، 6.3، 6.6، 6.5، 7.0، 7.1، 5.9، 6.4، 6.8، 5.8، 5.9، 6.7، 6.0،
5.9، 6.8، 6.9، 6.1، 6.3، 6.7، 6.3، 6.7، 5.8، 7.0، 7.1، 6.7، 5.9، 6.7،
6.4، 6.7، 5.9، 6.5، 6.1، 6.6، 6.6، 6.8، 7.0، 7.2، 5.8، 6.4، 6.5، 6.7،
6.6، 6.4، 6.7، 6.8، 6.0، 6.6 .

الحل (7-1):

لتنفيذ عملية الاختبار يتم الدخول الى البرنامج وفق الاجراءات التي تم سردها وتهيئة الملف (بالبيانات اعلاه) لاختباره للاختبار ثم اختيار الامر الرئيسي Analyze ومنه يتم اختيار الامر الفرعي Compare Means ومن الاخير يتم التأشير على One sample T test والكبس عليه فنحصل على مربع الحوار المبين في الشكل البياني (3-7)، وبعد القيام بتحويل المتغير الى المربع المختص الواقع على الجانب الايمن في مربع الحوار بواسطة الفأرة ثم نقله بواسطة السهم، واذا كنا ننوي الحصول على نتائج عند حدود ثقة تزيد على 95% يتم الكبس على ايقونة Option الواقعة في اسفل مربع الحوار الرئيسي، وبعد الكبس على ايقونة Option يظهر لنا مربع الحوار الملحق لتحديد درجة الثقة المطلوبة، اما اذا كانت درجة 95% وافية فلاحاجة الى مربع الحوار الاضافي، عندها كل ما نحتاجه هو الكبس على ايقونة Ok في مربع الحوار الرئيسي لنحصل على جدول المخرجات المبين في الجدول رقم (7-1) في ادناه :

الشكل البياني رقم (37)
 مربع الحوار الرئيسي One sample T test



جدول رقم (17)
مخرجات T-Test للمثال (17) للاختبار الاحادي

One- Sample Test

| | Test Value= 0 | | | | | |
|-----------------|---------------|----|----------------|---------------------|---|--------|
| | T | df | Sig.(2-tailed) | Mean Deifference | 95% Confidence Interval of the Difference | |
| | | | | | Lower | Upper |
| Sport Sample | 136.927 | 60 | .000 | 6.4967 | 6.4018 | 6.5916 |

ثانيا: تفسير مخرجات الاختبار الاحادي One sample T test

ومن المخرجات في الجدول رقم (7-1) نستدل على قبول ادعاء الشركة من ان متوسط مقاومة مضرب التنس هو 6.5 كغم وبانحراف معياري مقداره 0.45 كغم، حيث ان نتائج الاختبار مقبولة عند $\alpha = 0.000$ وهو يعتبر عالي المعنوية، وان المتوسط يقع تماما بين حدي الثقة 6.4018 و 6.5916 .

2. الاختبار في حالة عدم تساوي التباين (مجتمعين مستقلين)

Independent sample T-test

اولا: المفهوم والمدخلات

وهنا يعود المتوسطان لمجتمعين مختلفين، والهدف هو اختبار عما اذا كان الفرق بين متوسطي العينتين يعزى الى الصدفة او ان هذا الفرق جوهري. كمثال على ذلك اختبار مستوى جودة منتج يعود لشركتين مستقلتين، او اختبار نفس الظاهرة في بلدين مختلفين وهكذا .

مثال (7-2): تم جمع بيانات لعينتين من الاسر حجم كل منها 14 اسرة من مجتمعين مختلفين تتعلق بسؤال عن عمر الاطفال (بالاشهر) عند البدء بالمشي،

والمطلوب اختبار ان كان هناك فرق جوهري بين كلا المجتمعين في هذه الظاهرة.

$n_1 = 11.1, 10.5, 9.7, 13.4, 12.8, 10.0, 10.2, 9.2, 10.1, 9.5,$
 $10.1, 11.4, 12.4, 11.2$

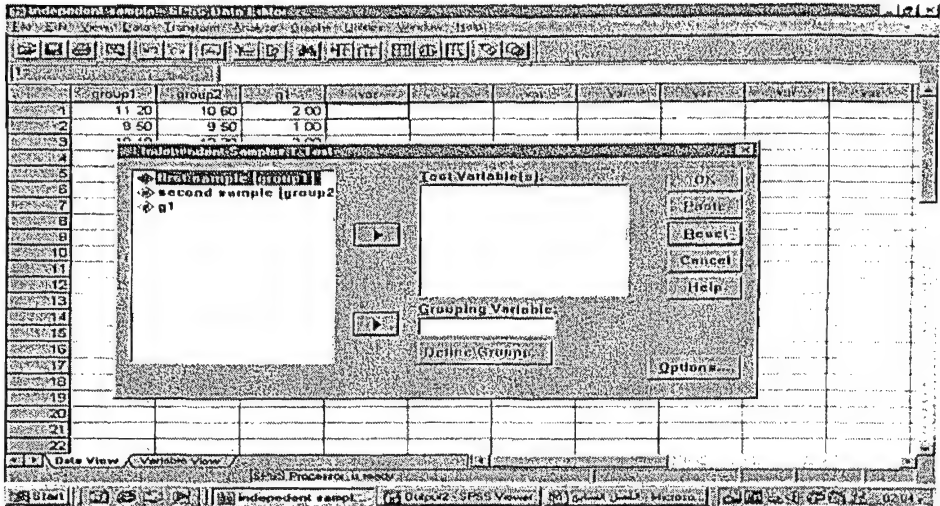
$n_2 = 10.2, 12.0, 13.2, 9.8, 9.6, 13.4, 12.6, 13.2, 12.3, 9.5,$
 $12.5, 9.8, 11.1, 10.6,$

الحل (2-7):

بعد تهيئة ملف البيانات الواردة في المثال نعين الامر الفرعي Compare Means ومنه يتم التأشير والكبس على اجراء Independent sample T-test لنحصل على مربع الحوار الرئيسي المبين في الشكل (4-7) لنقوم بنقل متغير العينتين اللتين تم وضعهما بنفس العمود على التوالي وتم تسميته (Group 1) وفي عمود المتغير الثاني تم اعطاء القيمة 1 امام المجموعة الاولى (العيينة الاولى) والقيمة 2 امام بيانات المجموعة الثانية (العيينة الثانية) وتم تسميته (Group 2) فنقوم بنقل متغير Group 1 الى الحقل الواقع على جهة اليمين بواسطة الفأرة ثم نقله بالسهم، ثم نقوم بنقل المتغير الآخر Group 2 الى الحقل الموجود في اسفل مربع الحوار الرئيسي والذي يحمل عنوان Grouping variable بعدها يتم الكبس على ايقونة Define variable للحصول على مربع الحوار الملحق المبين في الشكل (5-7) ليتم فيه تدوين 1 لمجموعة البيانات العينة الاولى و2 في الحقل الثاني للإشارة الى العينة الثانية وبعد الكبس على Continue نعود لمربع الحوار الرئيسي، وان لم نرغب بتغيير حدود الثقة عن 95%، تكون الخطوة اللاحقة هي الكبس على ايقونة Ok لنحصل على المخرجات المبينة في الجدول رقم (2-7)

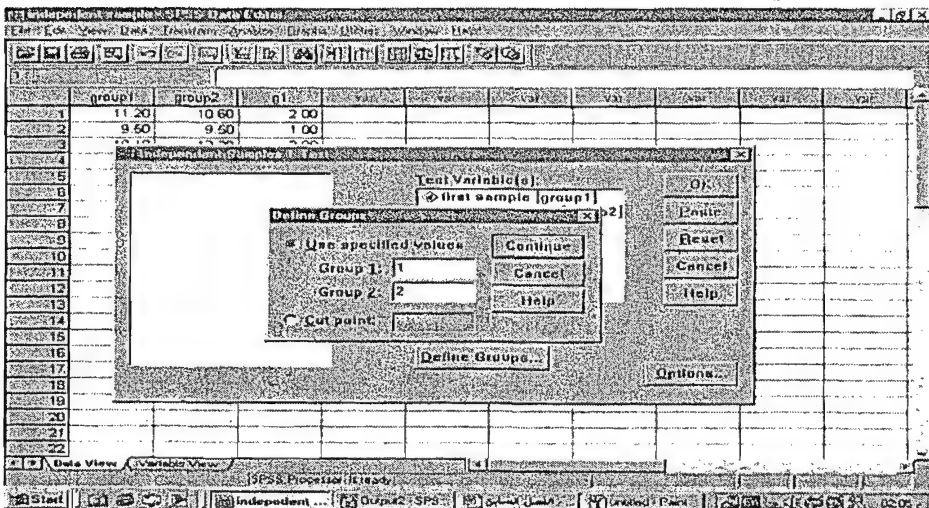
شكل بياني رقم (47)

مربع الحوار الرئيسي للاختبار عينتين مستقلتين Independent sample T-test



شكل بياني رقم (57)

مربع الحوار الملحق التعلق في Define Groups



جدول رقم (27)

مخرجات تحليل اختبار Independent sample T-test

Group Statistics

| | Second sample | N | Mean | Std. Deviation | Std. Error Mean |
|--------------|---------------|----|---------|----------------|-----------------|
| First Sample | 1.00 | 14 | 10.8286 | 1.2845 | .3433 |
| | 2.00 | 14 | 11.2929 | 1.3975 | .3735 |

Independent Samples Test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|--------------|--------------------------------|---|------|------------------------------|--------|--------------------|--------------------|--------------------------|---|-------|
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| First sample | Equal variances assumed | 1.360 | .254 | -.915 | 26 | .368 | -.4643 | .5073 | -1.5071 | .5785 |
| | Equal variances not assumed | | | -.915 | 25.817 | .369 | -.4643 | .5073 | -1.5074 | .5788 |

ثانياً: تفسير مخرجات اختبار عينتين مستقلتين Independent sample T-test

عند تعريف المجاميع في مرحلة Define Groups كنا قد اعطينا القيمة 1 لقيم العينة الاولى والقيمة 2 للعينة الثانية وبذلك فان الجدول الاول من المخرجات يبين بان متوسط المجموعة الاولى لمتغير العينة الاولى ومقداره 10.8288 ومتوسط المجموعة الثانية لنفس المتغير 11.2929 وكذلك الانحراف المعياري لكل من المجموعتين وعدد القيم ومقدار الخطأ المعياري كما يتبين في الجدول. اما الجزء الثاني من المخرجات الذي يضم نتائج الاختبار بافتراض تساوي التباينات والآخر هو عدم تساوي التباين والآخر هو المستهدف من عملية الاختبار بالنسبة لمثالنا، ومنه نستدل على وجود فروق جوهرية بين المجتمعين، حيث ان قيمة T المستخرجة من جانبين عند مستوى معنوية 0.369 (غير المعنوية). وان قيمة متوسط الفروق

0.5073 يقع خارج الحد الأدنى لحدود الثقة عند درجة 95% المبينة في آخر العمود من الجدول، وعليه نرفض فرضية التساوي بين المجتمعين .

3 اختبار T للمقارنات الزوجية Paired Samples T - test

اولا: المفهوم والمدخلات

ويتم استخدامه لقياس ظاهرة معينة بظروف مختلفة فمثلا لقياس نمو نباتات معينة عند تعرضها للشمس ونموها قبل تعرضها للشمس لاختبار ان كان هناك فرق جوهري في نموها بين كلا الحالتين. والفرضية التي يقوم عليها الاختبار هو ان المقارنة بين عينات غير مستقلة.

مثال (7-3): ادعى احد مكاتب الرشاقة بان نظام التدريب الذي لديه من شأنه ان يؤدي الى تخفيض جوهري في الوزن شهريا، ولاختبار صحة هذا الادعاء تم اختيار عينة حجمها 11 شخصا من الذين يسجلون لدى المكتب المذكور ودونت اوزانهم عند دخول الدورة، وكذلك بعد مرور شهر على التدريب، وكانت النتائج هي كما مبين في ادناه:

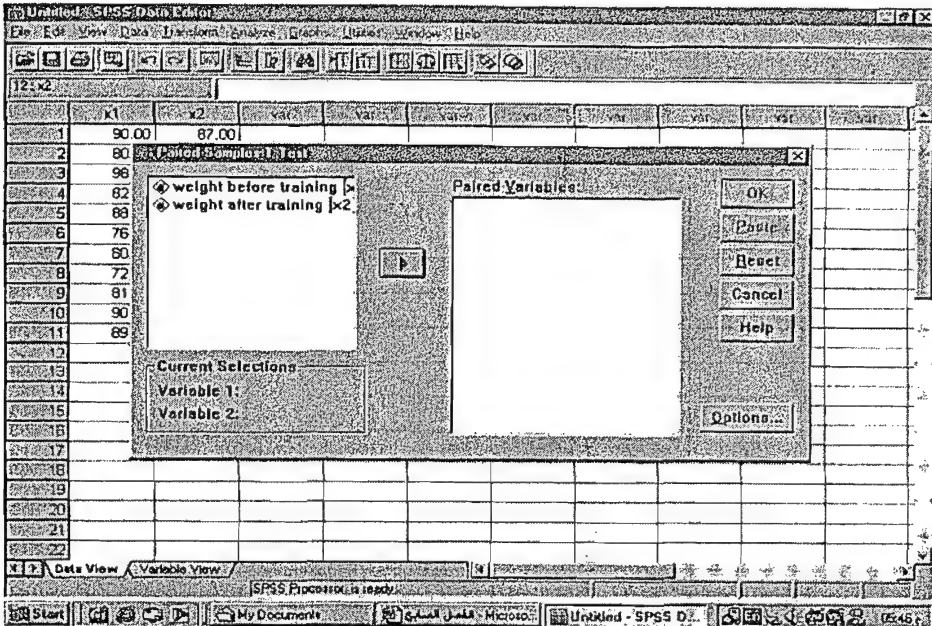
| الوزن بعد مرور شهر على الدورة (كغم) x2 | الوزن قبل دخول الدورة (كغم) x1 | التسلسل |
|---|-----------------------------------|---------|
| 87 | 90 | 1 |
| 78 | 80 | 2 |
| 94 | 98 | 3 |
| 80 | 82 | 4 |
| 86 | 88 | 5 |
| 73 | 76 | 6 |
| 79 | 80 | 7 |
| 70 | 72 | 8 |
| 79 | 81 | 9 |
| 85 | 90 | 10 |
| 86 | 89 | 11 |

الحل (3-7):

بعد تدوين البيانات في ملف، يتم اخضاعها للتحليل من خلال الامر الرئيسي Analyze ومن ثم اختيار الامر الفرعي Compare Means ومنه يتم الكبس على إجراء Paired Samples t- test فيظهر مربع الحوار الرئيسي المبين في الشكل البياني رقم (6-7)، فيتم تحويل المتغيرات الى الحقل الواقع الى اليمين باستخدام الفأرة ثم نقله بواسطة السهم. وبالكبس على ايقونة Ok نحصل على مخرجات التحليل المبينة في الجدول رقم (3-7) عند درجة ثقة مقدارها 95% .

شكل بياني رقم (6-7):

مربع الحوار الرئيسي لاختبار المقارنات الزوجية Paired Samples t - test



جدول رقم (37)

مخرجات تحليل اختبار المقارنات الزوجية Paired Samples t - test

Paired Samples Statistics

| | | Mean | N | Std. Deviation | Std. Error Mean |
|--------|------------------------|---------|----|----------------|-----------------|
| Pair 1 | Weight before training | 84.1818 | 11 | 7.4942 | 2.2596 |
| | Weight after training | 81.5455 | 11 | 6.8610 | 2.0687 |

Paired Samples Correlations

| | | N | Correlation | Sig. |
|--------|--|----|-------------|------|
| Pair 1 | Weight before training & Weight after training | 11 | .992 | .000 |

Paired Samples Test

| | | Paired Differences | | | | | t | df | Sig. (2-tailed) |
|--------|--|--------------------|----------------|-----------------|---|--------|-------|----|-----------------|
| | | Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | | | | |
| | | | | | Lower | Upper | | | |
| Pair 1 | Weight before training - Weight after training | 2.6364 | 1.1201 | .3377 | 1.8839 | 3.3888 | 7.807 | 10 | .000 |

ثانياً: تفسير مخرجات استخدام T-test للمقارنات الزوجية

Paired Samples t - test

بالرجوع الى جدول المخرجات رقم (3-7) اعلاه، نجد ان التحليل في مرحلته الاولى عرض متوسط وزن الاشخاص عند دخولهم للدورة وكان مقداره 84.182 كغم وانحراف معياري قدره 7.494 كغم، اصبحت بعد مرور شهر على التدريب

81.545 كغم وانحراف معياري مقداره 6.861 كغم، أي ان التحليل الأولي يشير الى تحقق انخفاض عام واضح في اوزان المشاركين في دورة التدريب، كما ان الانخفاض في حجم الانحراف المعياري يدل على ان هذا الانخفاض في الوزن اصبح يتسم بتجانس اكبر بعد مرور شهر على التدريب . ويعزز مخرجات المرحلة الاولى معامل الارتباط القوي بين قبل وبعد الدورة الذي بلغ 0.991 وهو مقبول عند معنوية 0.000 كما ان اشارة الارتباط جاءت باشارة موجبة مما يدل الى الدورة يمكن ان تحقق انخفاض اكبر مع ذوي الاوزان المرتفعة. ووفق هذه المؤشرات جاءت نتائج الاختبار لتثبت صحة ادعاء مركز الرشاقة وبمعنوية عالية عند $\alpha/2 = 0.000$.

4. اختبار مربعات كاي Chi-Square، χ^2

اولا: المفهوم والمخلات

يستخدم توزيع χ^2 لاختبار الفرضيات المتعلقة بالبيانات التي تكون على شكل توزيعات تكرارية، وتعتمد جميع اشكال استخدامه على اساس مقارنة التكرارات الحقيقية مع التكرارات المتوقعة وفقا لطبيعة التوزيع الاحتمالي للبيانات. ان الفرض الذي يقوم عليه الاختبار ان بيانات العينة مسحوبة من مجتمع طبيعي معلوم التباين. واهم مجالات استخدامه هي:

- حسن الجودة Goodness of Fit، أي لاختبار مدى التوافق بين القيم الحقيقية والفرضية للتكرارات. كما في حالة التوزيع العمري لعينة على فئات الاعمار مثلا.
- الاستقلالية Independency، لاختبار افتراض ان معايير التصنيف مع وحدات المجموعة هي مستقلة، كما في حالة تصنيف سكان ما حسب الحالة الاقتصادية او الاجتماعية، ومثل هذا الجدول يشار اليه بجدول التوافق Contingency Table.
- التجانس Consistency، أي مدى تجانس المجتمعات مع معايير التصنيف، كما في حالة تصنيف عينة من حوادث الطرق حسب نوع الحادث ونوع واسطة النقل.

مثال (7-4): اخذت عينة من حوادث الطرق في الاردن وتم تصنيفها حسب معياري نوع الحادث (اصطدام، دهس، انقلاب) ونوع واسطة النقل (صالون،

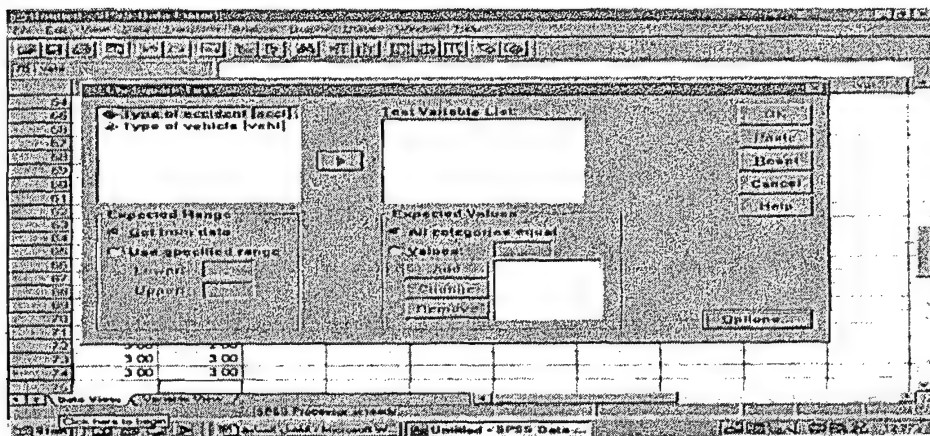
باص وبيك اب، لوري) وكما مبين في الجدول ادناه، والمطلوب اختبار ان كان هناك تجانس في نوع الحادث بين كافة انواع وسائط النقل المتورطة في هذه الحوادث .

| نوع واسطة النقل | نوع حادث الطريق | | | المجموع |
|-----------------|-----------------|-----|--------|---------|
| | اصطدام | دهس | انقلاب | |
| صالون | 20 | 16 | 4 | 40 |
| باص وبيك اب | 10 | 8 | 3 | 21 |
| لوري | 6 | 5 | 2 | 13 |
| المجموع | 36 | 29 | 9 | 74 |

الحل (4-7):

عقب اجراءات الدخول الى برنامج SPSS نقوم اولاً بتكوين ملف الحوادث وحسب الاجراءات التي تم شرحها في بداية الفصل الثالث، وتسجيل البيانات في مثل هذه الحالة يتم بتسمية متغيرين وليكن الاول لنوع الحادث وتدون فيه القيمة 1 للاصطدام و 2 للدهس و 3 للانقلاب، والمتغير الثاني لنوع واسطة النقل وتدون فيه القيمة 1 للصالون و 2 للباص والبيك اب و 3 للوري. وفي الخطوة الثانية يتم التأشير على الامر الرئيسي Analyze ومنه نختار الامر الفرعي Non Parametric test ليتم التأشير والكبس على طريقة Chi-Square ليظهر لنا مربع الحوار الرئيسي المبين في الشكل (7-7). نقوم بنقل المتغيرات الى الحقل المخصص على جهة اليمين باستخدام الفأرة والسهم، ونجد اسفل موقع المتغيرات حقلاً يشتمل على خيارين، الاول اختيار All categories equal او ان يقوم الباحث بتحديد نسب توقعاته من خلال اختيار ايقونة Values وهو ماتم استخدامه مع مثالنا في هذا الاختبار، روعي فيه نسب مايشكله كل نوع من الحوادث وكل نوع من وسائط النقل فاعطينا القيمة 4 اولاً لتخصص حوادث الاصطدام والقيمة 3 لحوادث الدهس والقيمة 1 لحوادث الانقلاب، وهذه القيم تنطبق على انواع وسائط النقل: صالون وباص وبيك اب ثم

شكل بياني رقم (77)
مربع الحوار الرئيسى لاختبار Chi-Square



جدول رقم (7-4)
مخرجات اختبار Chi-Square Test

| Type of accident | | | |
|------------------|------------|------------|----------|
| | Observed N | Expected N | Residual |
| 1.00 | 36 | 37.0 | -1.0 |
| 2.00 | 29 | 27.8 | 1.3 |
| 3.00 | 9 | 9.3 | -.3 |
| Total | 74 | | |

Type of vehicle

| | Observed N | Expected N | Residual |
|-------|------------|------------|----------|
| 1.00 | 40 | 37.0 | 3.0 |
| 2.00 | 26 | 27.8 | -1.8 |
| 3.00 | 8 | 9.3 | -1.3 |
| Total | 74 | | |

Type Statistics

| | Type of accident | Type of vehicle |
|--------------------------|------------------|-----------------|
| Chi- Square ^a | 0.090 | .523 |
| df | 2 | 2 |
| Asymp. Sig. | .956 | .770 |

ثانياً: تفسير مخرجات استخدام اختبار مربعات كاي χ^2 ؛ Chi-Square

ان النتائج الواردة في جدول المخرجات جاءت على اساس عدم تساوي تكرارات القيم المتوقعة، اي باستخدام خيار Values لكلا المتغيرين، ومنها نستدل ان قيمة χ^2 المحتسبة بالنسبة لنوع الحوادث المتورطة فيها وسائط النقل هي معنوية عند $0.044 = 1 - 0.956$ ومن ذلك نستدل بأن هناك تجانس في نوع الحادث وفقاً لمعيار نوع واسطة النقل، حيث القيمة الجدولية عند مستوى معنوية 0.05 ودرجات حرية 2 هي 0.103. بكلمة اخرى ان معيار وسائط النقل هي في تجانس مع نوع الحوادث، كما وان معيار حوادث الطرق متجانس مع انواع وسائط النقل، اما خيار Descriptive Statistics في حالة الرغبة باستخدامه يعطينا فكرة عن درجة التقارب في متوسطيهما وانحرافهما المعياري .

^a - 0 Cells (.0%) have expected frequencies less than 5. the minimum expected cell frequency is 9.3.

5- تحليل التباين بمعيار واحد One-Way Analysis of Variance

اولا: المفهوم والمدخلات

تناولت الاختبارات السابقة حالة متوسط عينة او عينتين من انها تعود الى مجتمع او مجتمعين لها نفس المعلمة، و تحليل التباين هو امتداد لاختبار T ليصبح بالامكان اختبار عينتين او اكثر مع تحليل طبيعة ومصدر التباين بين الظواهر المختلفة، حيث يقوم بتقسيم الاختلافات الكلية الى عدة اجزاء لتحديد مصدرها. ويقوم الاختبار على فرضية ان العينات تعود لمجتمعات موزعة طبيعيا وان عملية سحبها عشوائي وتبايناتها متساوية. وفي حالة عدم توفر الشرط الاخير يكون من المناسب استخدام اختبار بارتلليت Bartlet او هارتلي Hartly. ويعتمد الاختبار على مقياس F ونتائجه تنظم في جدول يدعى جدول تحليل التباين، ففي حالة التحليل بمعيار واحد يتم تصنيف قيم X الى K من المجاميع، فدرجات الطلبة تصنف حسب الشعب، وكل شعبة تضم n من الطلاب وعادة ما يشار اليها بالعناصر. ان الاختلاف بقيم X يعزى الى الاختلاف بين القيم الواقعة ضمن المجموعة الواحدة، والاختلاف بين المجاميع ذاتها، ونتبع في التالي اسلوب تحليل التباين بمعيار واحد وفي حالة تساوي حجوم المجاميع.

مثال (5-7): لاختيار عينة عشوائية، تم تقسيم مدينة ما الى اربعة مناطق، ومن كل منطقة تم اختيار عينة عشوائية تتكون من 9 مخازن لبيع المواد الغذائية، وكانت مبيعات كل مخزن اسبوعيا (بالآلاف الدنانير) كما هو مبين في الجدول التالي، والمطلوب معرفة ان كان هناك فرق معنوي في مبيعات مخازن المناطق الاربعة.

| المناطق | | | | المخازن |
|---------|----|----|----|---------|
| X4 | X3 | X2 | X1 | |
| 10 | 7 | 8 | 5 | 1 |
| 8 | 5 | 7 | 6 | 2 |
| 9 | 6 | 7 | 3 | 3 |
| 9 | 8 | 9 | 2 | 4 |
| 11 | 9 | 10 | 4 | 5 |
| 12 | 10 | 11 | 10 | 6 |
| 9 | 7 | 8 | 7 | 7 |
| 5 | 3 | 4 | 3 | 8 |
| 6 | 4 | 5 | 4 | 9 |

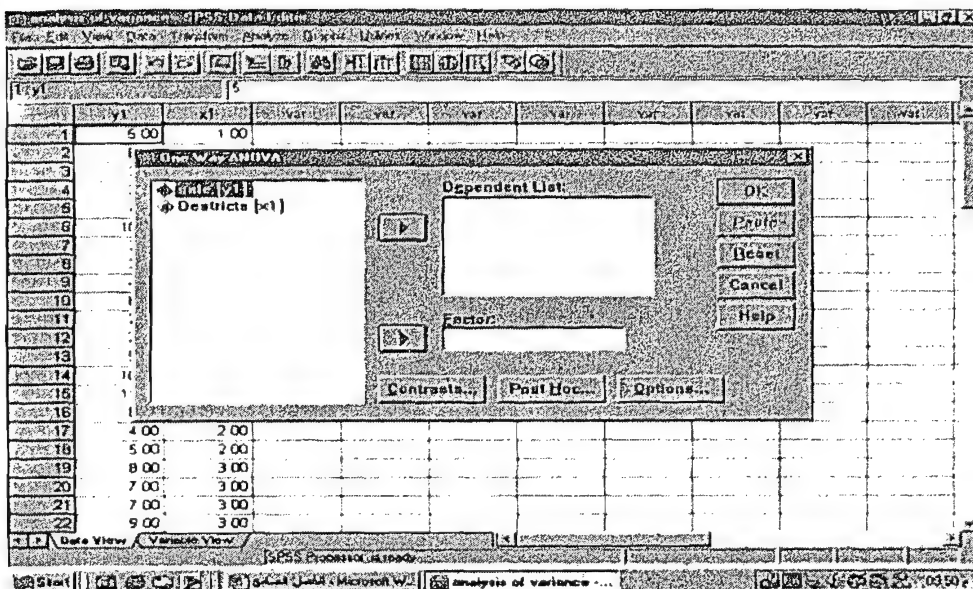
الحل (7-4):

وحيث لدينا متغير واحد بعدة مستويات (مجاميع)، عليه نستخدم في حل هذا المثال طريقة تحليل التباين بمعيار واحد One-Way Analysis of Variance، وعند متابعة الاجراءات المطلوبة في الحل سنحاول تجنب الدخول في المفاسل والتشعبات التي يمكن الحصول عليها ضمن مخرجات تحليل التباين، والاكتفاء بالاشارة اليها وذلك من اجل التبسيط والتشجيع على التعامل مع الامر من دون تعقيدات، خاصة وان هدف الكتاب هو الطلبة وغير المحترفين في مجال الاحصاء. كما سنحاول استخدام ذات المثال عند استخدام الطريقة اليدوية في الحل لاحقا للتأكد مما تؤول اليه نتائج المخرجات.

وكما هو السياق العام، فأول خطوة نحتاجها بعد الدخول الى برنامج SPSS هي اعداد ملف بالبيانات المطلوب تحليل تباينها، فيتم ادراج كافة بيانات المناطق في متغير واحد يطلق عليه لاغراض تحليل التباين بالمتغير التابع Dependent ليشمل قيم مبيعات كافة المخازن البالغ عددها 36 مخزناً، على ان نبدأ بمخازن المنطقة 1 ثم المنطقة 2 فالثالثة واخيرا المنطقة 4 على التوالي، بعدها يتم تكوين المتغير الآخر الذي يدعى هنا Factor (ويقصد به المتغير المستقل Independent) وتكون قيم هذا المتغير هي رموز المناطق على التوالي فيدرج رقم المنطقة امام قيم المخازن العائدة

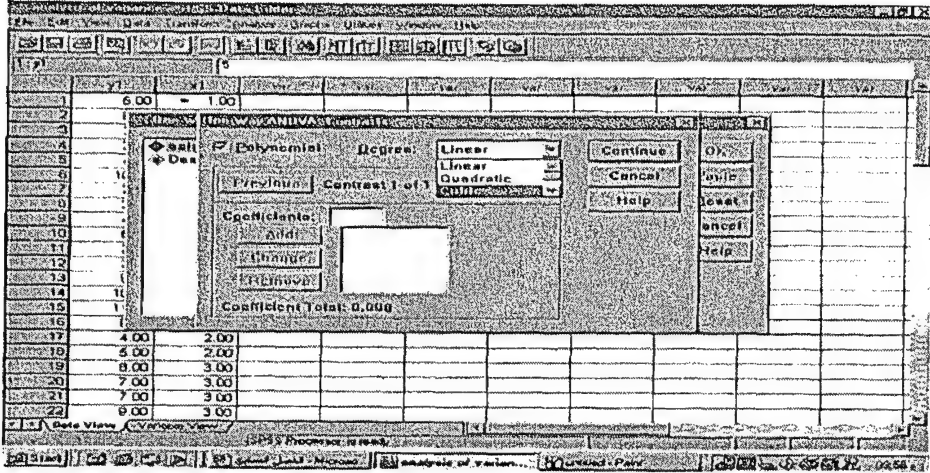
لها، لتأخذ الشكل الذي يظهر في الشكل البياني رقم (7-8). تليها الخطوة الثانية وهو استخدام الامر الرئيسي Analyze ومنه الامر الفرعي Compare Mean لنختار من الاخير One-Way Analysis of Variance فيظهر مربع الحوار الرئيسي المبين في الشكل رقم (7-8)، وفيه يتم نقل المتغير التابع الى الموقع المخصص له الى اليمين، والى اسفله المتغير الآخر الى حقل Factor.

الشكل البياني رقم (8-7) مربع الحوار الرئيسي لتحليل التباين Analysis of Variance



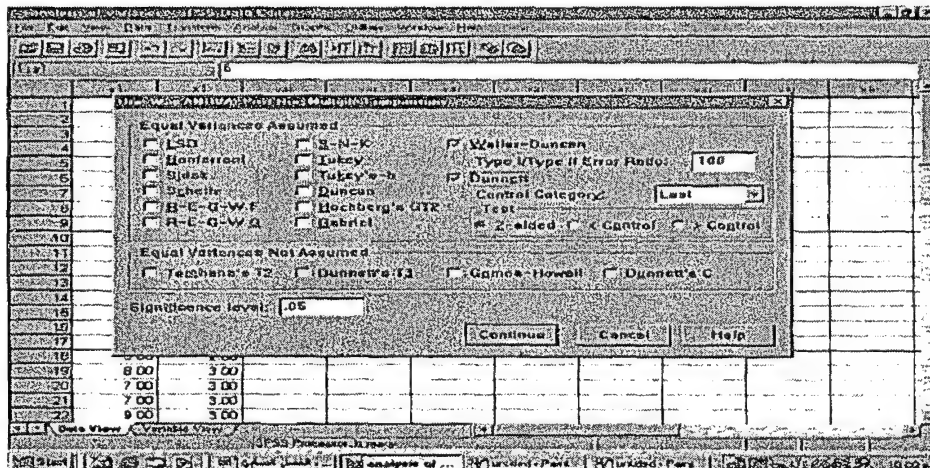
كما ونجد في اسفل مربع الحوار الرئيسي اعلاه الايقونات التالية:
 - Contrast ويتضمن مربع الحوار الملحق الطرق المعنوية بتقسيم مجموع المربعات بين المجاميع الى عناصر، ومن هذه الطرق ما يتعلق بالاتجاه الخطي او التربيعي او التكعيبي ... الخ، والشكل البياني رقم (7-9) يوضح الطرق المتوفرة في مربع الحوار الملحق.

شكل بياني رقم (97)
سريع الحوار الملحق Contrast



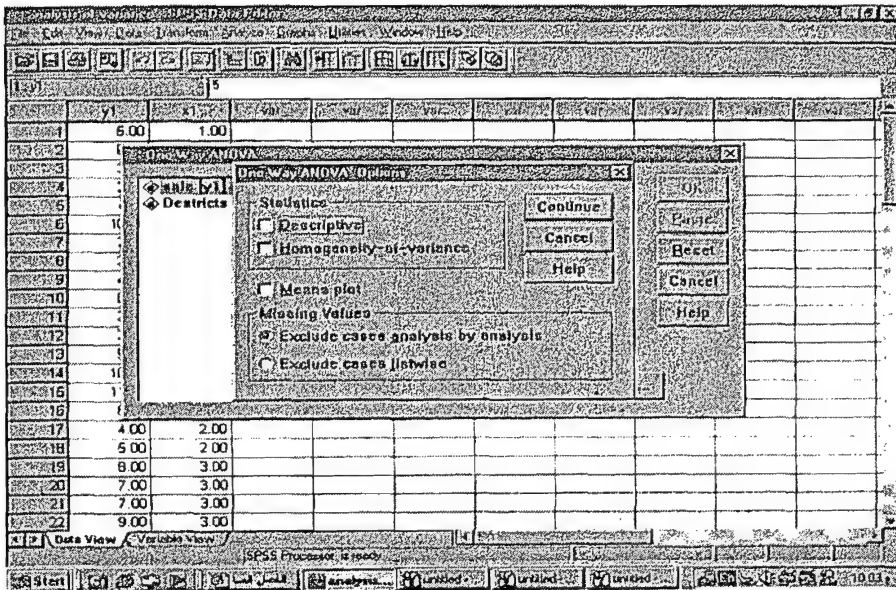
- Post Hoc Test ويستفاد منه في حالة وجود اختلافات واضحة بين متوسطات المجموعات المختلفة، وهناك خصائص محددة لكل من الخيارات العديدة المتوفرة في هذا المربع الحواري الملحق والمبين في الشكل البياني رقم (10-7).

شكل بياني رقم (107)
سريع الحوار الملحق Post Hoc Test



- Option وفي هذا المربع الملحق والمبين في الشكل البياني رقم (11-7) محوران، الاول هو Descriptive Statistics ويشمل مقاييس تتعلق بالمتوسطات والانحراف المعياري والخطأ المعياري والحد الأدنى والاعلى للقيم ودرجة الثقة. اما الثاني فهو Homogeneity of Variances وبواسطته يمكن استخدام إحصاءة Levene Static لفحص مدى تساوي تباين المجموعات من دون الحاجة الى فرضية التوزيع الطبيعي للبيانات .

شكل بياني رقم (11-7)
مربع الحوار الملحق Option



وباكتمال الاجراءات المطلوبة في مربع الحوار الرئيسي يتم الكبس على ايقونة Ok للحصول على المخرجات المجملّة المبينة في الجدول رقم (5-7) .

جدول رقم (7-5)

مخرجات تحليل التباين بمقيار واحد للمشال رقم (47) One-Way Analysis of Variance

Descriptives

Sale

| | | | | | 95% Confidence Interval for Mean | | | |
|-------|----|--------|----------------|------------|----------------------------------|-------------|---------|---------|
| | N | Mean | Std. Deviation | Std. Error | Lower Bound | Upper Bound | Minimum | Maximum |
| 1.00 | 9 | 4.8889 | 2.4721 | .8240 | 2.9887 | 6.7891 | 2.00 | 10.00 |
| 2.00 | 9 | 7.6667 | 2.2361 | .7454 | 5.9479 | 9.3855 | 4.00 | 11.00 |
| 3.00 | 9 | 7.6667 | 2.2361 | .7454 | 5.9479 | 9.3855 | 4.00 | 11.00 |
| 4.00 | 9 | 8.7778 | 2.2236 | .7412 | 7.0686 | 10.4870 | 5.00 | 12.00 |
| Total | 39 | 7.2500 | 2.6336 | .4389 | 6.3589 | 8.1411 | 2.00 | 12.00 |

ANOVA

Sale

| | Sum of Squares | df | Mean Square | F | Sig. |
|----------------|----------------|----|-------------|-------|------|
| Between Groups | 74.306 | 3 | 24.769 | 4.705 | .008 |
| Within Groups | 168.444 | 32 | 5.264 | | |
| Total | 242.750 | | | | |

ثانيا : تفسير مخرجات تحليل التباين بمقيار واحد

One-Way Analysis of Variance

عند التأمل في مخرجات الجزء المتعلق بمقاييس الاحصاء الوصفي نجد ان هناك فروقا بين متوسطات المبيعات للمناطق، مما انعكس ذلك نسبيا على حدي الثقة المبينة في العمود الاخير من الجدول ولكن ليس بدرجة جوهرية، وبالرجوع الى معنوية F عند 0.95 ثقة مع درجات 3 و 32 نجد ان القيمة الجدولية هي 2.89 مقابل 4.7 مما يستدل على رفض فرضية التجانس في مبيعات المناطق وقبول الفرضية البديلة القائلة بعدم التجانس. إلا أن قرار الرفض هذا يمكن تجنبه في حالة اجراء المقارنة عند مستوى معنوية أقل.

3.7. الطريقة اليدوية في اجراء الاختبارات وتحليل التباين

حيث قد تم التطرق الى مفاهيم كل من ادوات الاختبار وتحليل التباين والفرضيات التي تقوم عليها عند استخدام هذه الادوات مع برنامج SPSS، وتلافيا للتكرار فان الطريقة اليدوية هنا سنتناول الاجراءات التطبيقية، مع الاشارة الى ان نتائج اختبارات T-Test باستخدام الحاسوب ليس من الضروري ان تتساوى مع قيمة T المحسبة يدويا رغم التوصل الى نفس الاستنتاج من ناحية قرار الرفض او القبول، ويعود سبب ذلك الى انه في الحالة اليدوية نحتاج الى متوسط المجتمع μ وانحرافه σ وفي حالة عدم توفر معلمة الانحراف المعياري للمجتمع يتم الاستدلال عليها من الصيغة $\sigma = s/\sqrt{n}$ ، في حين يقوم الحاسوب بالاستدلال على هذه المعالم من بيانات العينة وفقا لحدود الثقة التي يتم استخراجها بموجب الصيغة:

$$\bar{X} - t_{\alpha/2} s/\sqrt{n} \leq \mu \leq \bar{X} + t_{\alpha/2} s/\sqrt{n} \text{ وكما يتبين من الامثلة التالية.}$$

1. الاختبار الاحادي One Sample T-test

ففي المثال (7-1) كان يتوفر وسط المجتمع وهو $\mu = 6.5$ و الانحراف المعياري ومقداره $\sigma = 0.45$ وبذلك يمكن تعويضها مباشرة في صيغة اختبار T الاحادي One Sample T-Test لنحصل على :

$$t = \frac{\bar{X} - \mu}{\sigma}$$

حيث ان: μ و \bar{X} هي متوسطا المجتمع والعينة على التوالي، σ الانحراف المعياري للمجتمع

$$t = \frac{6.4967 - 6.5}{0.45} = -0.0074$$

وحيث ان قيمة T المحسبة نقل كثيرا عن قيمة T الجدولية عند درجات حرية 60 والبالغة 1.994 مما يعني القبول بصحة ادعاء مصنع مضرب التنس بمعنوية عالية $\alpha = 0.000$ وهو نفس القرار الذي تم التوصل اليه في حالة استخدام الحاسوب.

2 الاختبار في حالة عينتين من مجتمعين مستقلين (مختلفي

التباين) Two- Independent Samples T test

وصيغة حسابها هو:

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{S_p^2/n_1 + S_p^2/n_2}}$$

حيث ان:

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

وبحل المثال (2-7) بدوياً، يكون لدينا:

$$n=14, s_2=1.3975, s_1=1.2845, \bar{X}=11.292, \bar{X}_1=10.0286$$

$$S_p^2 = 0.15123$$

$$T = \frac{1.2643}{0.147} = 8.6$$

وبمقارنة القيمة المحسوبة في اعلاه مع القيمة الجدولية عند مستوى معنوية ودرجات حرية 26 نجد ان $T_{\alpha/2, 26} = 2.766$. وبذلك نرفض الفرضية ونستدل على وجود فرق جوهري بين المجتمعين المسحوبة من العينات، وهي ذات الحصلة التي جاءت باستخدام الحاسوب مع برنامج SPSS.

3 اختبار المقارنات الزوجية Paired Samples T test

وصيغة حساب هذا الاختبار هو :

$$t = \frac{\bar{d} - \mu_d}{S_{\bar{d}}}$$

حيث ان d تشير الى الفروق بين قيم العينتين و $S_{\bar{d}} = s_d / \sqrt{n}$ بالرجوع الى المثال (3-7) يكون لدينا :

$$S_{\bar{d}} = s/\sqrt{n} = 1.12 / 3.1623 = 0.354 \quad , \quad \bar{d} = 2.63636$$

وبتطبيق صيغة الاختبار نحصل :

$$t = \frac{\bar{d} - \mu_d}{S_{\bar{d}}} = \frac{2.63636}{0.354} = 7.447$$

وكما نلاحظ فان قيمة t اعلاه المحتسبة بالطريقة اليدوية مقاربة جدا لتلك التي تم استخراجها بواسطة الحاسوب (ربما الفرق البسيط يعود إلى تقريب الكسور)، وهي اكبر من القيمة الجدولية عند درجات حرية 10 ومستوى دلالة 0.01، مما يترتب عليه رفض فرضية العدم القائلة بعدم وجود فرق جوهري بين الاشخاص قبل وبعد التدريب وقبول الفرضية البديلة عند مستوى معنوية 0.000 .

4. اختبار مربعات كاي Chi Square

وصيغة حسابه تأخذ الشكل التالي :

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

حيث ان: O_i تشير الى القيم الحقيقية، و $E_i = \frac{T_{cij}}{\sum C_{jj}}$ القيم الفرضية

C تشير الى الخلية، TC_{ij} حاصل ضرب مجموع عمود الخلية في مجموع صفها .

$\sum C_{jj}$ المجموع الكلي

وباجراء التطبيق على المثال (4-7) يكون لدينا:

| المجموع | نوع حادث الطريق | | | نوع واسطة النقل |
|---------|-----------------|-------------|-------------|-----------------|
| | انقلاب | دهس | اصطدام | |
| 40 | 4 (4.865) | 16 (15.676) | 20 (19.459) | صالون |
| 21 | 3 (2.554) | 8 (8.23) | 10 (10.216) | باص وبيك اب |
| 13 | 2 (1.581) | 5 (5.094) | 6 (6.324) | لوري |
| 74 | 9 | 29 | 36 | المجموع |

1. نبدأ أولا بإيجاد القيم الفرضية من خلال ضرب مجموع عمود الخلية بمجموع صفها مقسومة على المجموع الكلي فنحصل على القيم المحصورة بين قوسين في الجدول اعلاه .

2. نطبق صيغة مربعات كاي اعلاه فنحصل على :

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$= \frac{(20 - 19.459)^2}{19.459} + \frac{(16 - 15.676)^2}{15.676} + \frac{(4 - 4.865)^2}{4.865} + \frac{(10 - 10.216)^2}{10.216} + \frac{(8 - 8.23)^2}{8.23}$$

$$+ \frac{(3 - 2.554)^2}{2.554} + \frac{(6 - 6.324)^2}{6.324} + \frac{(5 - 5.094)^2}{5.094} + \frac{(2 - 1.581)^2}{1.581} = 0.3937$$

وعند مستوى معنوية 0.05 ودرجات حرية 4 نجد ان القيمة الجدولية هي 0.711 مما يعني ان القيمة المحتسبة تقل عن القيمة الجدولية، لنستدل على ان حوادث الطرق هي متجانسة وفقا لمعيار نوع واسطة النقل.

5- تحليل التباين بمعيار واحد One-Way Analysis of Variance

باجراء التحليل يدويا على ذات المثال (5-7) الذي تم اخضاعه لبرنامج SPSS فسنحتاج الى الخطوات التالية :

1. ايجاد القيم التالية:

$$\sum X_4 = 79 \quad \sum X_3 = 39 \quad \sum X_2 = 69 \quad \sum X_1 = 44, \quad \sum X_i = 251$$

$$\bar{X}_4 = 8.78 \quad \bar{X}_3 = 6.56 \quad \bar{X}_2 = 4.89 \quad \mu_{\bar{x}} = 6.97 \quad \sum X_i = 251$$

2. ستخرج مجموع مربعات الاختلاف بين المناطق (المجاميع) SSB

$$SSB = n \sum_{i=1}^k (\bar{X}_i - \mu_{\bar{x}})^2$$

$$= \{(4.89-6.97)^2 + (7.67-6.97)^2 + (6.56-6.97)^2 + (8.78-6.97)^2\} = 74.3454$$

3. نستخرج مجموع مربعات الاختلاف الكلي SST

$$SST = n \sum_{j=1}^9 \sum_{l=1}^4 (X_{lj} - \mu_{\bar{X}})^2$$

$$= (5-6.97) + (6-6.97) + \dots + (6-6.97) = 246.0876$$

4. نستخرج مجموع مربعات الاختلاف ضمن المناطق (المجاميع)

$$SSW = SST - SSB$$

$$= 246.087 - 74.3454 = 171.7422$$

5. وفي ضوء النتائج اعلاه يتم تنظيم جدول تحليل التباين كالآتي :

| F | متوسط المربعات | مجموع المربعات | درجات الحرية | مصدر التباين |
|--------|----------------|----------------|--------------|---------------|
| 4.6175 | 24.7818 | 74.3454 | k-1 = 3 | SSB |
| | 5.3669 | 171.7422 | k(n-1) = 32 | SSW |
| ----- | ----- | 246.0876 | nk-1 = 35 | المجموع الكلي |

6. وبمقارنة قيمة F المستخرجة في اعلاه مع القيمة الجدولية عند مستوى ثقة

0.95 ودرجات حرية 3 و32 والبالغة 2.89، نرفض فرضية التجانس بين

مبيعات المناطق، وقبولنا للفرضية البديلة. ومما يجدر الاشارة اليه هو

التقارب الشديد بين نتائج الاختبار المستخرجة باستخدام برنامج SPSS

والنتيجة المستخرجة يدويا في اعلاه .

تمارين الفصل السابع

تمرين (7-1): ادعى مقال أن كلفة بناء المتر المربع للسكن في بلدية A هي أعلى منها في بلدية B، فسحبت عينتان من المساكن في كل من البلديتين وتم حساب كلفة المتر المربع لكل منهما وكانت كما يلي :

عينة البلدية A: 32، 31، 30.5، 40، 34، 38، 37

عينة البلدية B: 32، 35، 37، 30.5، 38، 36، 35، 31، 31، 40، 31، 39

والمطلوب اختبار ادعاء المقال بالطريقة اليدوية عند مستوى معنوية 0.01، وباستخدام برنامج SPSS .

تمرين (7-2): لدينا عينة تتكون من 10 نباتات منزلية تم قياس نموها لمدة شهر وهي في الظل، ثم تم تعريضها لضوء أكثر لنفس الفترة وتم قياس نموها بعد مرور شهر أيضاً كما هو مبين في الجدول التالي. والمطلوب اختبار أن كان هناك فرق جوهري بين أطوال نموها قبل وبعد التعرض للضوء الإضافي، مستخدماً برنامج SPSS ومن ثم بالطريقة اليدوية .

| أطوال النباتات (سم) | | التسلسل |
|--------------------------|--------------------------|---------|
| بعد التعرض للضوء الإضافي | قبل التعرض للضوء الإضافي | |
| 33 | 31 | 1 |
| 32 | 33 | 2 |
| 36 | 35 | 3 |
| 29 | 30 | 4 |
| 39 | 36 | 5 |
| 38 | 37 | 6 |
| 41 | 41 | 7 |
| 40 | 35 | 8 |
| 43 | 39 | 9 |
| 34 | 32 | 10 |

تمرين (3-7): ا- المطلوب اختبار وعند مستوى معنوية $\alpha = 0.01$ باستخدام الطريقة اليدوية ان كان معياري تصنيف الدخل وتصنيف السكن حسب المساحة، مستقلة عن بعضها لعينة من الاسر عددها $n = 465$ المبينة في الجدول التالي .

ب- استخدم الاعداد المبينة في الجدول لتنظيم ملف بالمتغيرين المذكورين بما يتناسب واخضاعها للتحليل باستخدام برنامج SPSS.

ج- اخضاع الملف الذي يتم تنظيمه لاجراء اختبار Chi-Square .

تمرين (4-7): تم استخدام 4 طرق مع 4 مجاميع من الطلبة لتعليمهم جدول الضرب، وكانت النتائج كما هو مبين في الجدول ادناه، والمطلوب اختبار ان كانت هناك فروق جوهرية بين الطرق الاربع عند مستوى معنوية $\alpha = 0.05$ ، باستخدام كل من برنامج SPSS والطريقة اليدوية.

| مجاميع الطلبة | | | | طريقة التعليم |
|---------------|----|---|---|---------------|
| 4 | 3 | 2 | 1 | |
| 9 | 8 | 6 | 7 | 1 |
| 8 | 10 | 9 | 8 | 2 |
| 6 | 10 | 8 | 7 | 3 |
| 9 | 5 | 6 | 8 | 4 |

تمرين (5-7): استعملت ثلاثة انواع من الاطارات على العجلات الخلفية لنفس النوع من وسائط النقل وعلى ثلاثة انواع من الطرق، والبيانات في الجدول التالي تمثل عدد الكيلومترات (بالآف) التي قطعت قبل انتهاء صلاحية الاطارات. والمطلوب اختبار ان كان معدل المسافة المقطوعة من قبل الاطارات المختلفة على الطرق المختلفة متساوياً عند مستوى معنوية 0.01.

| انواع الطرق | | | نوع الاطار |
|-------------|------|------|------------|
| C | B | A | |
| 9.7 | 8.2 | 10.3 | 1 |
| 8.6 | 7.7 | 11.2 | |
| 8.3 | 7.9 | 9.8 | |
| 9.0 | 7.2 | 10.1 | |
| 8.7 | 10.1 | 11.1 | 2 |
| 8.4 | 9.8 | 10.8 | |
| 8.9 | 9.8 | 11.5 | |
| 8.0 | 10.3 | 11.6 | |
| 6.7 | 8.7 | 7.7 | 3 |
| 6.9 | 8.5 | 7.4 | |
| 7.4 | 9.4 | 7.5 | |
| 7.4 | 8.8 | 8.0 | |

الملاحق الإحصائية

□ ملحق رقم (12)

جدول الأرقام العشوائية Random Digits

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 87024 | 74221 | 69721 | 44518 | 58804 | 04860 | 18127 | 16855 | 61558 | 15430 |
| 04852 | 03436 | 72753 | 99836 | 37513 | 91341 | 53517 | 92094 | 54386 | 44563 |
| 33592 | 45845 | 52015 | 72030 | 23071 | 92933 | 84219 | 39455 | 57792 | 14216 |
| 68121 | 53688 | 56812 | 34869 | 28573 | 51079 | 94677 | 23993 | 88241 | 97735 |
| 25062 | 10428 | 43930 | 69033 | 73395 | 83469 | 25990 | 12971 | 73728 | 03856 |
| 78183 | 44396 | 11064 | 92153 | 96293 | 00825 | 21079 | 78337 | 19739 | 13684 |
| 70209 | 23316 | 32828 | 00927 | 61841 | 64754 | 91125 | 01206 | 06691 | 50868 |
| 94342 | 91040 | 94035 | 02650 | 36284 | 91162 | 07950 | 36178 | 42536 | 49869 |
| 92503 | 29854 | 24116 | 61149 | 49266 | 82303 | 54924 | 58251 | 23928 | 20703 |
| 71646 | 57503 | 82416 | 22657 | 72359 | 30085 | 13037 | 39608 | 77439 | 49318 |
| 51809 | 70780 | 41544 | 27828 | 84321 | 07714 | 25865 | 97896 | 01924 | 62028 |
| 88504 | 21620 | 07292 | 71021 | 80929 | 45042 | 08703 | 45894 | 24521 | 49942 |
| 33186 | 49273 | 87542 | 41086 | 29615 | 81101 | 43707 | 87031 | 36101 | 15137 |
| 40068 | 35043 | 05280 | 62921 | 30122 | 65119 | 40512 | 26855 | 40842 | 83244 |
| 76401 | 68461 | 20711 | 12007 | 19209 | 28259 | 49820 | 76415 | 51534 | 63574 |
| 47014 | 93729 | 74235 | 47808 | 52473 | 03145 | 92563 | 05837 | 70023 | 33169 |
| 67147 | 48017 | 90741 | 53647 | 55007 | 36607 | 29360 | 83163 | 79024 | 26155 |
| 86987 | 62924 | 93157 | 70947 | 07336 | 49541 | 81386 | 26968 | 38311 | 99885 |
| 58973 | 47026 | 78574 | 08804 | 22960 | 32850 | 67944 | 92303 | 61216 | 72948 |
| 71635 | 86749 | 40369 | 94639 | 40731 | 54012 | 03972 | 98581 | 45604 | 34885 |
| 60971 | 54212 | 32596 | 03052 | 84150 | 36798 | 62635 | 26210 | 95685 | 87089 |
| 06599 | 60910 | 66315 | 96690 | 19039 | 39878 | 44688 | 65146 | 02482 | 73130 |
| 89960 | 27162 | 66264 | 71024 | 18708 | 77974 | 40473 | 87155 | 35834 | 03114 |
| 03930 | 56898 | 61900 | 44036 | 90012 | 17673 | 54167 | 82396 | 39468 | 49566 |
| 31338 | 28729 | 02095 | 07429 | 35718 | 86882 | 37513 | 51560 | 08872 | 33717 |
| 29782 | 33287 | 27400 | 42915 | 49914 | 68221 | 56088 | 06112 | 95481 | 30094 |
| 68493 | 88796 | 94771 | 89418 | 62045 | 40681 | 15941 | 05962 | 44378 | 64349 |
| 42534 | 31925 | 94158 | 90197 | 62874 | 53659 | 33433 | 48610 | 14698 | 54761 |
| 76126 | 41049 | 43363 | 52461 | 00552 | 93352 | 58497 | 16347 | 87145 | 73668 |
| 80434 | 73037 | 69008 | 36801 | 25520 | 14161 | 32300 | 04187 | 80668 | 07499 |
| 81301 | 39731 | 53857 | 19690 | 39998 | 49829 | 12399 | 70867 | 44498 | 17385 |
| 54521 | 42350 | 82908 | 51212 | 70208 | 39891 | 64871 | 67448 | 42988 | 32600 |
| 82530 | 22869 | 87276 | 06678 | 36873 | 61198 | 87748 | 07531 | 29592 | 39612 |
| 81338 | 64309 | 45798 | 42954 | 95565 | 02789 | 83017 | 82936 | 67117 | 17709 |
| 58264 | 60374 | 32610 | 17879 | 96900 | 68029 | 06993 | 84288 | 35401 | 56317 |
| 77023 | 46829 | 21332 | 77383 | 15547 | 29332 | 77698 | 89878 | 20489 | 71800 |
| 29750 | 59902 | 78110 | 59018 | 87548 | 10225 | 15774 | 70778 | 56086 | 08117 |
| 08288 | 38411 | 69886 | 64918 | 29055 | 87607 | 37452 | 38174 | 31431 | 46173 |
| 93908 | 94810 | 22057 | 94240 | 89918 | 16561 | 92716 | 66461 | 22337 | 64718 |
| 06341 | 25883 | 42574 | 80202 | 57287 | 95120 | 69332 | 19036 | 43326 | 98697 |
| 23240 | 94741 | 55622 | 79479 | 34606 | 51079 | 09476 | 10695 | 49618 | 63037 |
| 96370 | 19171 | 40441 | 05002 | 33165 | 28693 | 45027 | 73791 | 23047 | 32976 |
| 97050 | 16194 | 61095 | 26533 | 81738 | 77032 | 60551 | 31605 | 95212 | 81078 |
| 40833 | 12169 | 10712 | 78345 | 48236 | 45086 | 61654 | 94929 | 69169 | 70561 |
| 95676 | 13582 | 25664 | 60838 | 88071 | 50052 | 63188 | 50346 | 65618 | 17517 |
| 28030 | 14185 | 13226 | 99566 | 45483 | 10079 | 22945 | 23903 | 11695 | 10694 |
| 60202 | 32586 | 87466 | 83357 | 95516 | 31258 | 66309 | 40615 | 30572 | 60842 |
| 46530 | 48755 | 02308 | 79508 | 53422 | 50805 | 08896 | 06963 | 93922 | 99423 |
| 53151 | 95839 | 01745 | 46462 | 81463 | 28669 | 60179 | 17880 | 75875 | 34562 |
| 80272 | 64398 | 88249 | 06792 | 98424 | 66842 | 49129 | 98939 | 34173 | 49883 |

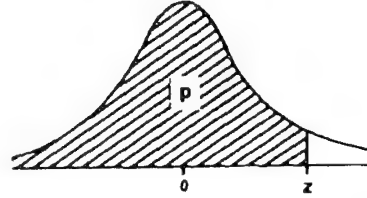
ملحق رقم (15)

جدول نسب التوزيع الطبيعي Percentage of Normal Distribution

The table gives the values of z satisfying

$$P(Z \leq z) = p$$

where Z is a normally distributed random variable with zero mean and unit variance.



| p | .00 | .01 | .02 | .03 | .04 | .05 | .06 | .07 | .08 | .09 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.50 | 0.000 | 0.025 | 0.050 | 0.075 | 0.100 | 0.126 | 0.151 | 0.176 | 0.202 | 0.228 |
| 0.60 | 0.253 | 0.279 | 0.305 | 0.332 | 0.358 | 0.385 | 0.412 | 0.440 | 0.468 | 0.496 |
| 0.70 | 0.524 | 0.553 | 0.583 | 0.613 | 0.643 | 0.674 | 0.706 | 0.739 | 0.772 | 0.806 |
| 0.80 | 0.842 | 0.878 | 0.915 | 0.954 | 0.994 | 1.036 | 1.080 | 1.126 | 1.175 | 1.227 |
| 0.90 | 1.282 | 1.341 | 1.405 | 1.476 | 1.555 | | | | | |

| p | .000 | .001 | .002 | .003 | .004 | .005 | .006 | .007 | .008 | .009 |
|------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.95 | 1.645 | 1.655 | 1.665 | 1.675 | 1.685 | 1.695 | 1.706 | 1.717 | 1.728 | 1.739 |
| 0.96 | 1.751 | 1.762 | 1.774 | 1.787 | 1.799 | 1.812 | 1.825 | 1.838 | 1.852 | 1.866 |
| 0.97 | 1.881 | 1.896 | 1.911 | 1.927 | 1.943 | 1.960 | 1.977 | 1.995 | 2.014 | 2.034 |
| 0.98 | 2.054 | 2.075 | 2.097 | 2.120 | 2.144 | 2.170 | 2.197 | 2.226 | 2.257 | 2.290 |
| 0.99 | 2.326 | 2.366 | 2.409 | 2.457 | 2.512 | 2.576 | 2.652 | 2.748 | 2.878 | 3.090 |

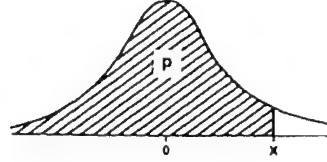
ملحق رقم (25)

نسب توزیع قییم (T) Percentage of the Student's t-Distribution

The table gives the values of x satisfying

$$P(X \leq x) = p$$

where X is a random variable having the Student's t -distribution with ν degrees of freedom.



| ν \ P | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 |
|-----------|-------|-------|--------|--------|--------|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 |

| ν \ P | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 |
|-----------|-------|-------|-------|-------|-------|
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 |
| 31 | 1.309 | 1.696 | 2.040 | 2.453 | 2.744 |
| 32 | 1.309 | 1.694 | 2.037 | 2.449 | 2.738 |
| 33 | 1.308 | 1.692 | 2.035 | 2.445 | 2.733 |
| 34 | 1.307 | 1.691 | 2.032 | 2.441 | 2.728 |
| 35 | 1.306 | 1.690 | 2.030 | 2.438 | 2.724 |
| 36 | 1.306 | 1.688 | 2.028 | 2.434 | 2.719 |
| 37 | 1.305 | 1.687 | 2.026 | 2.431 | 2.715 |
| 38 | 1.304 | 1.686 | 2.024 | 2.429 | 2.712 |
| 39 | 1.304 | 1.685 | 2.023 | 2.426 | 2.708 |
| 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 |
| 45 | 1.301 | 1.679 | 2.014 | 2.412 | 2.690 |
| 50 | 1.299 | 1.676 | 2.009 | 2.403 | 2.678 |
| 55 | 1.297 | 1.673 | 2.004 | 2.396 | 2.668 |
| 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 |
| 65 | 1.295 | 1.669 | 1.997 | 2.385 | 2.654 |
| 70 | 1.294 | 1.667 | 1.994 | 2.381 | 2.648 |
| 75 | 1.293 | 1.665 | 1.992 | 2.377 | 2.643 |
| 80 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 |
| 85 | 1.292 | 1.663 | 1.988 | 2.371 | 2.635 |
| 90 | 1.291 | 1.662 | 1.987 | 2.368 | 2.632 |
| 95 | 1.291 | 1.661 | 1.985 | 2.366 | 2.629 |
| 100 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 |
| 125 | 1.288 | 1.657 | 1.979 | 2.357 | 2.616 |
| 150 | 1.287 | 1.655 | 1.976 | 2.351 | 2.609 |
| 200 | 1.286 | 1.653 | 1.972 | 2.345 | 2.601 |
| ∞ | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 |

ملحق رقم (35)

Critical Values of Correlation Coefficient قيم معامل الارتباط البسيطة

| One tail Two tail n | 10% 20% | 5% 10% | 2.5% 5% | 1% 2% | 0.5% 1% |
|---------------------------|------------|-----------|------------|----------|------------|
| 4 | 0.8000 | 0.9000 | 0.9500 | 0.9800 | 0.9900 |
| 5 | 0.6870 | 0.8054 | 0.8783 | 0.9343 | 0.9587 |
| 6 | 0.6084 | 0.7293 | 0.8114 | 0.8822 | 0.9172 |
| 7 | 0.5509 | 0.6694 | 0.7545 | 0.8329 | 0.8745 |
| 8 | 0.5067 | 0.6215 | 0.7067 | 0.7887 | 0.8343 |
| 9 | 0.4716 | 0.5822 | 0.6664 | 0.7498 | 0.7977 |
| 10 | 0.4428 | 0.5494 | 0.6319 | 0.7155 | 0.7646 |
| 11 | 0.4187 | 0.5214 | 0.6021 | 0.6851 | 0.7348 |
| 12 | 0.3981 | 0.4973 | 0.5760 | 0.6581 | 0.7079 |
| 13 | 0.3802 | 0.4762 | 0.5529 | 0.6339 | 0.6835 |
| 14 | 0.3646 | 0.4575 | 0.5324 | 0.6120 | 0.6614 |
| 15 | 0.3507 | 0.4409 | 0.5140 | 0.5923 | 0.6411 |
| 16 | 0.3383 | 0.4259 | 0.4973 | 0.5742 | 0.6226 |
| 17 | 0.3271 | 0.4124 | 0.4821 | 0.5577 | 0.6055 |
| 18 | 0.3170 | 0.4000 | 0.4683 | 0.5425 | 0.5897 |
| 19 | 0.3077 | 0.3887 | 0.4555 | 0.5285 | 0.5751 |
| 20 | 0.2992 | 0.3783 | 0.4438 | 0.5155 | 0.5614 |
| 21 | 0.2914 | 0.3687 | 0.4329 | 0.5034 | 0.5487 |
| 22 | 0.2841 | 0.3598 | 0.4227 | 0.4921 | 0.5368 |
| 23 | 0.2774 | 0.3515 | 0.4132 | 0.4815 | 0.5256 |
| 24 | 0.2711 | 0.3438 | 0.4044 | 0.4716 | 0.5151 |
| 25 | 0.2653 | 0.3365 | 0.3961 | 0.4622 | 0.5052 |
| 26 | 0.2598 | 0.3297 | 0.3882 | 0.4534 | 0.4958 |
| 27 | 0.2546 | 0.3233 | 0.3809 | 0.4451 | 0.4869 |
| 28 | 0.2497 | 0.3172 | 0.3739 | 0.4372 | 0.4785 |
| 29 | 0.2451 | 0.3115 | 0.3673 | 0.4297 | 0.4705 |
| 30 | 0.2407 | 0.3061 | 0.3610 | 0.4226 | 0.4629 |
| 31 | 0.2366 | 0.3009 | 0.3550 | 0.4158 | 0.4556 |
| 32 | 0.2327 | 0.2960 | 0.3494 | 0.4093 | 0.4487 |
| 33 | 0.2289 | 0.2913 | 0.3440 | 0.4032 | 0.4421 |
| 34 | 0.2254 | 0.2869 | 0.3388 | 0.3972 | 0.4357 |
| 35 | 0.2220 | 0.2826 | 0.3338 | 0.3916 | 0.4296 |
| 36 | 0.2187 | 0.2785 | 0.3291 | 0.3862 | 0.4238 |
| 37 | 0.2156 | 0.2746 | 0.3246 | 0.3810 | 0.4182 |
| 38 | 0.2126 | 0.2709 | 0.3202 | 0.3760 | 0.4128 |
| 39 | 0.2097 | 0.2673 | 0.3160 | 0.3712 | 0.4076 |
| 40 | 0.2070 | 0.2638 | 0.3120 | 0.3665 | 0.4026 |
| 41 | 0.2043 | 0.2605 | 0.3081 | 0.3621 | 0.3978 |
| 42 | 0.2018 | 0.2573 | 0.3044 | 0.3578 | 0.3932 |
| 43 | 0.1993 | 0.2542 | 0.3008 | 0.3536 | 0.3887 |
| 44 | 0.1970 | 0.2512 | 0.2973 | 0.3496 | 0.3843 |
| 45 | 0.1947 | 0.2483 | 0.2940 | 0.3457 | 0.3801 |
| 46 | 0.1925 | 0.2455 | 0.2907 | 0.3420 | 0.3761 |
| 47 | 0.1903 | 0.2429 | 0.2876 | 0.3384 | 0.3721 |
| 48 | 0.1883 | 0.2403 | 0.2845 | 0.3348 | 0.3683 |
| 49 | 0.1863 | 0.2377 | 0.2816 | 0.3314 | 0.3646 |
| 50 | 0.1843 | 0.2353 | 0.2787 | 0.3281 | 0.3610 |
| 60 | 0.1678 | 0.2144 | 0.2542 | 0.2997 | 0.3301 |
| 70 | 0.1550 | 0.1982 | 0.2352 | 0.2776 | 0.3060 |
| 80 | 0.1448 | 0.1852 | 0.2199 | 0.2597 | 0.2864 |
| 90 | 0.1364 | 0.1745 | 0.2072 | 0.2449 | 0.2702 |
| 100 | 0.1292 | 0.1654 | 0.1966 | 0.2324 | 0.2565 |

ملحق رقم (45)

نسب توزيع (F) Percentage of F-Distribution

The table below corresponds to $p=0.995$ and should be used for one-tail tests at significance level 0.5% or two-tail tests at significance level 1%.

| α | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 | 20 | 25 | 30 | 40 | 50 | 100 | α |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| 1 | 161.4 | 199.5 | 215.7 | 227.9 | 237.9 | 246.7 | 254.3 | 261.9 | 269.4 | 276.9 | 284.4 | 291.9 | 300.0 | 308.0 | 315.9 | 323.9 | 332.0 | 340.0 | 348.0 | 1 |
| 2 | 199.5 | 246.7 | 269.4 | 284.4 | 291.9 | 299.4 | 306.9 | 314.4 | 321.9 | 329.4 | 336.9 | 344.4 | 352.0 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 2 |
| 3 | 215.7 | 269.4 | 291.9 | 306.9 | 314.4 | 321.9 | 329.4 | 336.9 | 344.4 | 352.0 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 3 |
| 4 | 227.9 | 284.4 | 306.9 | 321.9 | 329.4 | 336.9 | 344.4 | 352.0 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 4 |
| 5 | 237.9 | 291.9 | 314.4 | 329.4 | 336.9 | 344.4 | 352.0 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 5 |
| 6 | 246.7 | 306.9 | 329.4 | 344.4 | 352.0 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 6 |
| 7 | 254.3 | 314.4 | 336.9 | 352.0 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 7 |
| 8 | 261.9 | 321.9 | 344.4 | 359.5 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 8 |
| 9 | 269.4 | 329.4 | 352.0 | 367.0 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 9 |
| 10 | 276.9 | 336.9 | 359.5 | 374.5 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 10 |
| 11 | 284.4 | 344.4 | 367.0 | 382.0 | 390.0 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 11 |
| 12 | 291.9 | 352.0 | 374.5 | 389.5 | 398.0 | 406.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 12 |
| 15 | 308.0 | 374.5 | 398.0 | 414.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 518.0 | 526.0 | 534.0 | 15 |
| 20 | 315.9 | 382.0 | 406.0 | 422.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 518.0 | 526.0 | 534.0 | 542.0 | 20 |
| 25 | 323.9 | 390.0 | 414.0 | 430.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 518.0 | 526.0 | 534.0 | 542.0 | 550.0 | 25 |
| 30 | 332.0 | 398.0 | 422.0 | 438.0 | 446.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 518.0 | 526.0 | 534.0 | 542.0 | 550.0 | 558.0 | 30 |
| 40 | 348.0 | 414.0 | 438.0 | 454.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 518.0 | 526.0 | 534.0 | 542.0 | 550.0 | 558.0 | 566.0 | 574.0 | 40 |
| 50 | 360.0 | 422.0 | 446.0 | 462.0 | 470.0 | 478.0 | 486.0 | 494.0 | 502.0 | 510.0 | 518.0 | 526.0 | 534.0 | 542.0 | 550.0 | 558.0 | 566.0 | 574.0 | 582.0 | 50 |
| 100 | 398.0 | 462.0 | 494.0 | 510.0 | 518.0 | 526.0 | 534.0 | 542.0 | 550.0 | 558.0 | 566.0 | 574.0 | 582.0 | 590.0 | 598.0 | 606.0 | 614.0 | 622.0 | 630.0 | 100 |
| α | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | α |

The table below corresponds to $p=0.99$ and should be used for one-tail tests at significance level 1% or two-tail tests at significance level 2%.

| α | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 15 | 20 | 25 | 30 | 40 | 50 | 100 | α |
|----------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| 1 | 405.2 | 500.0 | 540.3 | 562.5 | 576.4 | 585.9 | 592.8 | 598.1 | 602.2 | 605.6 | 608.3 | 610.6 | 612.8 | 614.9 | 616.9 | 618.8 | 620.6 | 622.4 | 624.1 | 1 |
| 2 | 499.6 | 594.0 | 634.3 | 656.5 | 670.4 | 679.9 | 686.8 | 692.1 | 696.2 | 700.0 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 2 |
| 3 | 540.3 | 634.3 | 656.5 | 670.4 | 679.9 | 686.8 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 3 |
| 4 | 562.5 | 656.5 | 670.4 | 679.9 | 686.8 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 4 |
| 5 | 576.4 | 670.4 | 679.9 | 686.8 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 5 |
| 6 | 585.9 | 679.9 | 686.8 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 6 |
| 7 | 592.8 | 686.8 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 7 |
| 8 | 598.1 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 8 |
| 9 | 598.1 | 692.1 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 9 |
| 10 | 602.2 | 696.2 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 10 |
| 11 | 605.6 | 699.9 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 11 |
| 12 | 608.3 | 703.7 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 12 |
| 15 | 612.8 | 707.3 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 15 |
| 20 | 614.9 | 710.8 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 760.7 | 20 |
| 25 | 616.9 | 714.2 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 760.7 | 763.5 | 25 |
| 30 | 618.8 | 717.5 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 760.7 | 763.5 | 766.3 | 30 |
| 40 | 620.6 | 720.7 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 760.7 | 763.5 | 766.3 | 769.1 | 40 |
| 50 | 622.4 | 723.9 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 760.7 | 763.5 | 766.3 | 769.1 | 771.9 | 50 |
| 100 | 624.1 | 727.0 | 729.9 | 732.7 | 735.5 | 738.3 | 741.1 | 743.9 | 746.7 | 749.5 | 752.3 | 755.1 | 757.9 | 760.7 | 763.5 | 766.3 | 769.1 | 771.9 | 774.7 | 100 |
| α | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | 7.679 | α |

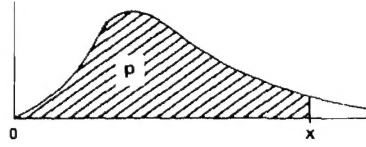
ملحق رقم (17)

نسب توزيع مربعات كاي χ^2 - Distribution Percentage of the χ^2

The table gives the values of x satisfying

$$P(X \leq x) = p$$

where X is a χ^2 random variable with v degrees of freedom.



| $v \backslash p$ | 0.005 | 0.01 | 0.025 | 0.05 | 0.1 | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 |
|------------------|---------|--------|--------|--------|--------|---------|---------|---------|---------|---------|
| 1 | 0.00004 | 0.0002 | 0.001 | 0.004 | 0.016 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 |
| 2 | 0.010 | 0.020 | 0.051 | 0.103 | 0.211 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 |
| 3 | 0.072 | 0.115 | 0.216 | 0.352 | 0.584 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 |
| 4 | 0.207 | 0.297 | 0.484 | 0.711 | 1.064 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 |
| 5 | 0.412 | 0.554 | 0.831 | 1.145 | 1.610 | 9.236 | 11.070 | 12.833 | 15.086 | 16.750 |
| 6 | 0.676 | 0.872 | 1.237 | 1.635 | 2.204 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 |
| 7 | 0.989 | 1.239 | 1.690 | 2.167 | 2.893 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 |
| 8 | 1.344 | 1.646 | 2.180 | 2.733 | 3.490 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 |
| 9 | 1.735 | 2.088 | 2.700 | 3.325 | 4.168 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 |
| 10 | 2.156 | 2.558 | 3.247 | 3.940 | 4.865 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 |
| 11 | 2.603 | 3.053 | 3.816 | 4.575 | 5.578 | 17.275 | 19.675 | 21.920 | 24.725 | 26.757 |
| 12 | 3.074 | 3.571 | 4.404 | 5.226 | 6.304 | 18.549 | 21.026 | 23.337 | 26.217 | 28.306 |
| 13 | 3.565 | 4.107 | 5.009 | 5.892 | 7.042 | 19.812 | 22.362 | 24.736 | 27.688 | 29.819 |
| 14 | 4.075 | 4.660 | 5.629 | 6.571 | 7.790 | 21.064 | 23.685 | 26.119 | 29.141 | 31.319 |
| 15 | 4.601 | 5.229 | 6.262 | 7.261 | 8.547 | 22.307 | 24.996 | 27.488 | 30.578 | 32.801 |
| 16 | 5.142 | 5.812 | 6.908 | 7.962 | 9.312 | 23.542 | 26.296 | 28.845 | 32.000 | 34.267 |
| 17 | 5.697 | 6.406 | 7.564 | 8.672 | 10.085 | 24.769 | 27.587 | 30.191 | 33.409 | 35.718 |
| 18 | 6.265 | 7.015 | 8.231 | 9.390 | 10.865 | 25.989 | 28.869 | 31.526 | 34.805 | 37.156 |
| 19 | 6.844 | 7.633 | 8.907 | 10.117 | 11.651 | 27.204 | 30.144 | 32.852 | 36.191 | 38.582 |
| 20 | 7.434 | 8.260 | 9.591 | 10.851 | 12.443 | 28.412 | 31.410 | 34.170 | 37.566 | 39.997 |
| 21 | 8.034 | 8.897 | 10.283 | 11.591 | 13.240 | 29.615 | 32.671 | 35.479 | 38.932 | 41.401 |
| 22 | 8.643 | 9.542 | 10.982 | 12.338 | 14.041 | 30.813 | 33.924 | 36.781 | 40.289 | 42.796 |
| 23 | 9.260 | 10.196 | 11.689 | 13.091 | 14.848 | 32.007 | 35.172 | 38.076 | 41.638 | 44.181 |
| 24 | 9.886 | 10.856 | 12.401 | 13.848 | 15.659 | 33.196 | 36.415 | 39.364 | 42.980 | 45.559 |
| 25 | 10.520 | 11.524 | 13.120 | 14.611 | 16.473 | 34.382 | 37.652 | 40.646 | 44.314 | 46.928 |
| 26 | 11.160 | 12.198 | 13.844 | 15.379 | 17.292 | 35.563 | 38.885 | 41.923 | 45.642 | 48.290 |
| 27 | 11.808 | 12.879 | 14.573 | 16.151 | 18.114 | 36.741 | 40.113 | 43.195 | 46.963 | 49.645 |
| 28 | 12.461 | 13.565 | 15.308 | 16.928 | 18.939 | 37.916 | 41.337 | 44.461 | 48.270 | 50.993 |
| 29 | 13.121 | 14.256 | 16.047 | 17.708 | 19.768 | 39.087 | 42.557 | 45.722 | 49.580 | 52.336 |
| 30 | 13.787 | 14.953 | 16.791 | 18.493 | 20.599 | 40.256 | 43.773 | 46.979 | 50.892 | 53.672 |
| 31 | 14.458 | 15.655 | 17.539 | 19.281 | 21.434 | 41.422 | 44.985 | 48.232 | 52.191 | 55.003 |
| 32 | 15.134 | 16.362 | 18.291 | 20.072 | 22.271 | 42.585 | 46.194 | 49.480 | 53.486 | 56.328 |
| 33 | 15.815 | 17.074 | 19.047 | 20.867 | 23.110 | 43.745 | 47.400 | 50.725 | 54.776 | 57.648 |
| 34 | 16.501 | 17.789 | 19.806 | 21.664 | 23.952 | 44.903 | 48.602 | 51.966 | 56.061 | 58.964 |
| 35 | 17.192 | 18.509 | 20.569 | 22.465 | 24.797 | 46.059 | 49.802 | 53.203 | 57.342 | 60.275 |
| 36 | 17.887 | 19.233 | 21.336 | 23.269 | 25.643 | 47.212 | 50.998 | 54.437 | 58.619 | 61.581 |
| 37 | 18.586 | 19.960 | 22.106 | 24.075 | 26.492 | 48.363 | 52.192 | 55.668 | 59.892 | 62.883 |
| 38 | 19.289 | 20.691 | 22.878 | 24.884 | 27.343 | 49.513 | 53.384 | 56.896 | 61.162 | 64.181 |
| 39 | 19.996 | 21.426 | 23.654 | 25.695 | 28.196 | 50.660 | 54.572 | 58.120 | 62.428 | 65.476 |
| 40 | 20.707 | 22.164 | 24.433 | 26.509 | 29.051 | 51.805 | 55.758 | 59.342 | 63.691 | 66.766 |
| 45 | 24.311 | 25.901 | 28.366 | 30.612 | 33.350 | 57.505 | 61.656 | 65.410 | 69.957 | 73.166 |
| 50 | 27.991 | 29.707 | 32.357 | 34.764 | 37.689 | 63.167 | 67.505 | 71.420 | 76.154 | 79.490 |
| 55 | 31.735 | 33.570 | 36.398 | 38.958 | 42.060 | 68.796 | 73.311 | 77.380 | 82.292 | 85.749 |
| 60 | 35.534 | 37.485 | 40.482 | 43.188 | 46.459 | 74.397 | 79.082 | 83.298 | 88.379 | 91.952 |
| 65 | 39.383 | 41.444 | 44.603 | 47.450 | 50.883 | 79.973 | 84.821 | 89.177 | 94.422 | 98.105 |
| 70 | 43.275 | 45.442 | 48.758 | 51.739 | 55.329 | 85.527 | 90.531 | 95.023 | 100.425 | 104.215 |
| 75 | 47.206 | 49.475 | 52.942 | 56.054 | 59.795 | 91.061 | 96.217 | 100.839 | 106.393 | 110.286 |
| 80 | 51.172 | 53.540 | 57.153 | 60.391 | 64.278 | 96.578 | 101.879 | 106.629 | 112.329 | 116.321 |
| 85 | 55.170 | 57.634 | 61.389 | 64.749 | 68.777 | 102.079 | 107.522 | 112.393 | 118.236 | 122.325 |
| 90 | 59.196 | 61.754 | 65.647 | 69.126 | 73.291 | 107.565 | 113.145 | 118.136 | 124.116 | 128.299 |
| 95 | 63.250 | 65.898 | 69.925 | 73.520 | 77.818 | 113.038 | 118.752 | 123.858 | 129.973 | 134.247 |
| 100 | 67.328 | 70.065 | 74.222 | 77.929 | 82.358 | 118.498 | 124.342 | 129.561 | 135.807 | 140.169 |

قائمة المصادر

1. د. عبد الحميد عبد المجيد البلداوي، الاحصاء للعلوم الادارية والتطبيقية، دار الشروق للنشر والتوزيع، عمان - الاردن، 1997
2. د. عبد الحميد عبد المجيد البلداوي، طرق المعاينة التطبيقية، جامعة السابع من ابريل، الزاوية - ليبيا، 1995
3. رانية عثمان المشاركة، برنامج التحليل الاحصائي SPSS، مكتبة الراتب العلمية، عمان - الاردن، 1997
4. د. محمد ازهر السماك - د. قبيس الفهادي - صفاء الصفواوي، اصول البحث العلمي، جامعة صلاح الدين - العراق، 1986

5. Draper N. and Smith H., Applied Regression Analysis, John Wiley and Sons Inc. , London , 1990
6. Snedecor G. and Cochran G. Statistical Methods, 7th Edition, The Iowa State University Press , U.S.A. , 1980

المركز الإسلامي الثقافي

مكتبة سماحة آية الله العظمى

السيد محمد حسين فضل الله العامة

الرقم 5096 /

يستهدف هذا الكتاب توفير صورة متكاملة للباحث والدارس عما يحتاجه لإعداد بحث أو دراسة ابتداءً من مرحلة التخطيط للبحث وماهية المواضيع اللازم تناولها في عملية التخطيط ، ومن ثم كيفية تحديد عدد الاستبيانات أو حجم البيانات والمعلومات المطلوب توفيرها للبحث وكذلك سرد أنواع العينات التي يمكن اعتمادها في اختيار وحدات العينة .

ويتناول الكتاب بعد مرحلة جمع وتوفير البيانات عملية التحليل الإحصائي باستخدام برنامج SPSS مع التفصيل في كيفية استخدام البرنامج وكيفية اختيار الأسلوب أو الأدوات الإحصائية التي تتوافق مع أهداف البحث والدراسة ، مع التطرق إلى تفسير مخرجات التحليل الناتجة من توظيف البرنامج ، بالإضافة إلى التحليل يدوياً للوقوف على الأسس والمعادلات التي تم استخدامها في برنامج SPSS وتناولت المرحلة الأخيرة كافة خطوات ومستلزمات التحليل بما في ذلك الأساليب الإحصائية المتقدمة كاختبار الفروض وتحليل التباين والانحدار وتحليل المركبات (العوامل) وغيرها .



دار الشروق للنشر والتوزيع

المركز الرئيسي-عمان- الأردن/ تلفون 4618190-4618191-4624321

فاكس 4610065 ص ب: 926463 عمان 11118 الأردن

فروع الجامعة الأردنية / تلفون 5358352

E.mail: shorokjo@nol.com.jo

www.shorok.com

ISBN 9957-00-318-6

